

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2005, 17:39:11 ; Search time 89.0148 Seconds
(without alignments)
373.662 Million cell updates/sec

Title: US-10-054-873-4
Perfect score: 463
Sequence: 1 FVNQHLCGSHLVEALYLVCG.....IVEQCCTSICSILYQLENYCN 86

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------|--------------------|
| 1 | 463 | 100.0 | 86 | 1 | AAP40829 | Aap40829 Sequence |
| 2 | 463 | 100.0 | 86 | 2 | AAR84061 | Aar84061 Human ins |
| 3 | 463 | 100.0 | 86 | 2 | AAY42858 | Aay42858 Human ins |
| 4 | 463 | 100.0 | 86 | 3 | AAB12770 | Aab12770 Human pro |
| 5 | 463 | 100.0 | 86 | 5 | AAM48218 | Aam48218 Human pro |
| 6 | 463 | 100.0 | 86 | 7 | ADC64463 | Adc64463 Amino aci |
| 7 | 463 | 100.0 | 86 | 7 | ADF16632 | Adf16632 Human alb |
| 8 | 463 | 100.0 | 86 | 7 | ADH21860 | Adh21860 Human lon |
| 9 | 463 | 100.0 | 86 | 8 | ADT93277 | Adt93277 Human nat |

| | | | | | | | |
|----|-----|-------|-----|---|----------|----------|-----------|
| 10 | 463 | 100.0 | 87 | 1 | AAP20036 | Aap20036 | Human pro |
| 11 | 463 | 100.0 | 87 | 1 | AAP40217 | Aap40217 | Sequence |
| 12 | 463 | 100.0 | 87 | 1 | AAP50127 | Aap50127 | Sequence |
| 13 | 463 | 100.0 | 87 | 1 | AAP50060 | Aap50060 | Synthetic |
| 14 | 463 | 100.0 | 87 | 1 | AAP61090 | Aap61090 | Sequence |
| 15 | 463 | 100.0 | 87 | 2 | AAR32367 | Aar32367 | Proinsuli |
| 16 | 463 | 100.0 | 88 | 2 | AAR07682 | Aar07682 | Modified |
| 17 | 463 | 100.0 | 88 | 2 | AAR33855 | Aar33855 | hpI. 3/20 |
| 18 | 463 | 100.0 | 92 | 2 | AAR20467 | Aar20467 | Yeast alp |
| 19 | 463 | 100.0 | 92 | 8 | ADL24442 | Adl24442 | Modified |
| 20 | 463 | 100.0 | 93 | 1 | AAP90102 | Aap90102 | Synthetic |
| 21 | 463 | 100.0 | 96 | 2 | AA08004 | Aay08004 | Human pro |
| 22 | 463 | 100.0 | 96 | 2 | AA017830 | Aao17830 | Human pro |
| 23 | 463 | 100.0 | 96 | 8 | ADL24437 | Adl24437 | Modified |
| 24 | 463 | 100.0 | 97 | 2 | AAR68898 | Aar68898 | Human pro |
| 25 | 463 | 100.0 | 97 | 3 | AAB12773 | Aab12773 | Human pro |
| 26 | 463 | 100.0 | 110 | 1 | AAP10042 | Aap10042 | Sequence |
| 27 | 463 | 100.0 | 110 | 1 | AAP10053 | Aap10053 | Sequence |
| 28 | 463 | 100.0 | 110 | 1 | AAP40309 | Aap40309 | Sequence |
| 29 | 463 | 100.0 | 110 | 2 | AA06608 | Aay06608 | Human pre |
| 30 | 463 | 100.0 | 110 | 2 | ADF77582 | Adf77582 | Human ins |
| 31 | 463 | 100.0 | 110 | 3 | AA044367 | Aay44367 | Human pro |
| 32 | 463 | 100.0 | 110 | 3 | AA070366 | Aay70366 | Human Ins |
| 33 | 463 | 100.0 | 110 | 3 | AAB26765 | Aab26765 | Human pre |
| 34 | 463 | 100.0 | 110 | 3 | AAB06144 | Aab06144 | Human ins |
| 35 | 463 | 100.0 | 110 | 4 | AAE10337 | Aae10337 | Human pre |
| 36 | 463 | 100.0 | 110 | 4 | AAB35424 | Aab35424 | Secretory |
| 37 | 463 | 100.0 | 110 | 4 | AAG65677 | Aag65677 | Human pro |
| 38 | 463 | 100.0 | 110 | 5 | ABG60634 | Abg60634 | Human ins |
| 39 | 463 | 100.0 | 110 | 5 | ABG31590 | Abg31590 | Human pre |
| 40 | 463 | 100.0 | 110 | 6 | ABR55862 | Abr55862 | Humanised |
| 41 | 463 | 100.0 | 110 | 6 | ADA09218 | Ada09218 | Human Pre |
| 42 | 463 | 100.0 | 110 | 7 | ADC51569 | Adc51569 | Human pro |
| 43 | 463 | 100.0 | 110 | 7 | ADE56710 | Ade56710 | Human Pro |
| 44 | 463 | 100.0 | 110 | 7 | ADD46938 | Add46938 | Human Pro |
| 45 | 463 | 100.0 | 110 | 7 | ADE57650 | Ade57650 | Human Pro |

ALIGNMENTS

RESULT 1

AAP40829

ID AAP40829 standard; protein; 86 AA.

XX

AC AAP40829;

XX

DT 09-SEP-2004 (revised)

DT 25-MAR-2003 (revised)

DT 03-AUG-1992 (first entry)

XX

DE Sequence of human insulin precursor.

XX

KW Insulin precursor; connecting peptide; diabetes; hormone.

XX

OS Homo sapiens.

OS Unidentified.

XX
 FH Key Location/Qualifiers
 FT Region 1. .30
 FT /label= chain B
 FT Modified-site 1
 FT /label= F-NH2-R
 FT /note= "H or a chemically or enzymatically cleavable AA
 FT residue or peptide residue"
 FT Disulfide-bond 7. .72
 FT Disulfide-bond 19. .85
 FT Peptide 31. .65
 FT /label= connecting peptide
 FT Region 66. .86
 FT /label= chain A
 FT Disulfide-bond 71. .76
 FT Modified-site 86
 FT /label= N-OH
 XX
 PN US4430266-A.
 XX
 PD 07-FEB-1984.
 XX
 PF 16-FEB-1982; 82US-00349397.
 XX
 PR 27-MAR-1980; 80US-00134389.
 PR 28-NOV-1980; 80US-00210696.
 XX
 PA (ELIL) LILLY & CO ELI.
 XX
 PI Frank BH;
 XX
 DR WPI; 1984-049032/08.
 XX
 PT Insulin precursor prodn. from linear S-sulphonate and mercaptan - in
 PT single step without separate oxidn.
 XX
 PS Claim 17; Col 4; 8pp; English.
 XX
 CC The inventors claim a method for the prepn. of an insulin precursor in
 CC which the A-chain and B-chain are joined through a connecting peptide.
 CC The connecting peptide joins the A-chain at the amino group of A-1 to the
 CC B-chain at the carboxyl group of B-30. The method is pref. for the prepn.
 CC of human insulin precursor (see AAP40829). The SQs of the connecting
 CC peptides of a number of species are given (see AAP40828, AAP40830-39).
 CC (Updated on 25-MAR-2003 to correct PA field.)
 CC
 CC Revised record issued on 09-SEP-2004 : Correction to Feature Table Key
 XX
 SQ Sequence 86 AA;

Query Match 100.0%; Score 463; DB 1; Length 86;
 Best Local Similarity 100.0%; Pred. No. 1.5e-43;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60

```
Qy      61 SLQKRGIVEQCCTSICSLYQLENYCN 86
      |||
Db      61 SLQKRGIVEQCCTSICSLYQLENYCN 86
```

RESULT 2

AAR84061

ID AAR84061 standard; protein; 86 AA.

XX

AC AAR84061;

XX

DT 22-AUG-1996 (first entry)

XX

DE Human insulin.

XX

KW Insulin; transformation; gene expression; fungi; fungal cell; hormone;

KW A-chain; C-chain; glycosylation.

XX

OS Homo sapiens.

XX

| FH | Key | Location/Qualifiers |
|----|-----|---------------------|
|----|-----|---------------------|

| | | |
|----|-----|---------|
| FT | CDS | 1. .261 |
|----|-----|---------|

```
FT          /*tag=  a
```

```
FT /product= "Insulin."
```

XX

PN EP704527-A2.

XX

PD 03-APR-1996.

XX

PF 03-AUG-1995; 95EP-00112210.

XX

PR 05-AUG-1994; 94HR-00000432.

XX

PA (PLIV) PLIVA PHARM & CHEM FAB.

XX

PI Mestric S, Punt PJ, Valinger R, Van Den Hondel CAMJJ;

XX

DR WPI; 1996-129917/18.

DR N-PSDB; AAT17830, AAT17831.

XX

PT DNA encoding human insulin precursors - which comprise B- and A-chains

PT linked via amino acid chain contg. 1 or more glycosylation sites, for

PT prepn. of insulin in fungal cells.

XX

PS Disclosure; Fig 1; 32pp; English.

XX

CC DNA sequences encoding insulin precursors of formula B-Pg-A, where B and
CC A represent B- and A-chains of insulin respectively, and Pg represents a
CC modified C-peptide or any number of amino acids comprising at least one
CC glycosylation consensus site, can be inserted into expression vectors
CC which in turn can be used to transform fungal host cells. The fungal
CC cells are then cultured and the insulin expressed in such cells can be
CC harvested

XX

SQ Sequence 86 AA;

Query Match 100.0%; Score 463; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.5e-43;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
        |||
Db      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
        |||

Qy      61 SLQKRGIVEQCCTSICSLYQLENYCN 86
        |||
Db      61 SLQKRGIVEQCCTSICSLYQLENYCN 86

```

RESULT 3

AAV42858

ID AAY42858 standard; protein; 86 AA.

XX

AC AAY42858:

XX

DT 19-JAN-2000 (first entry)

XX

DE Human insulin precursor, SEQ ID 5.

XX

KW Insulin; precursor; growth hormone; chaperone; intramolecular; folding;
KW conformation; chimeric protein; cleavable; recombinant; production;
KW yield.

XX

OS Homo sapiens.

XX

PN WO9950302-A1.

XX

PD 07-OCT-1999.

XX

PF 31-MAR-1998; 98WO-CN000052.

XX

PR 31-MAR-1998; 98WO-CN000052.

XX

PA (TONG-) TONGHUA GANTECH BIOTECHNOLOGY LTD.

XX

PI Gan Z;

XX

DR WPI; 1999-610839/52.

XX

PT New chimeric proteins containing human growth hormone fragment, used
PT particularly for the production of human insulin.

XX

PS Claim 10; Page 29; 46pp; English.

XX

CC This sequence represents a human insulin precursor comprising insulin A
CC and B chains separated by a 34 residue peptide sequence. This insulin
CC precursor can be a component of chimeric proteins which additionally
CC contains an N-terminal fragment of human growth hormone (hGH) and a
CC cleavable peptide linker (AAY42857). The hGH portion of the chimeric
CC protein acts as an intramolecular chaperone (IMC) for the insulin
CC precursor, enabling it to fold correctly. The cleavable peptide linker
CC has a C-terminal Arg residue which enables the hGH portion of the
CC chimeric protein to be removed after folding has taken place. Production

CC of recombinant human insulin via an hGH-proinsulin chimeric protein can
 CC provide human insulin with correctly linked cysteine bridges with fewer
 CC necessary procedural steps, and hence resulting in a higher yield of
 CC human insulin. The IMC sequences not only protect insulin sequences from
 CC intracellular degradation by a microorganism host, but also promote the
 CC folding of the fused insulin precursor, facilitate the solubility of the
 CC fusion protein and decrease the intermolecular interactions among the
 CC fusion proteins, thus allowing folding of the fused insulin precursor at
 CC commercially useful high concentrations. The procedural steps of cyanogen
 CC bromide cleavage, oxidative sulphytolysis and related purification steps
 CC can thus be eliminated, along with the use of high concentrations of
 CC mercaptan or the use of hydrophobic absorbent resins

XX

SQ Sequence 86 AA;

Query Match 100.0%; Score 463; DB 2; Length 86;
 Best Local Similarity 100.0%; Pred. No. 1.5e-43;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
 |||

Db 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60

Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 86
 |||

Db 61 SLQKRGIVEQCCTSICSLYQLENYCN 86

RESULT 4

AAB12770

ID AAB12770 standard; protein; 86 AA.

XX

AC AAB12770;

XX

DT 22-NOV-2000 (first entry)

XX

DE Human proinsulin protein sequence SEQ ID NO:2.

XX

KW Human; insulin-like growth factor 1; IGF-1; proinsulin; insulin; mutant;
 KW variant; insulin-like growth factor binding protein; IGFBP-1; IGFBP-3;
 KW antidiabetic; neuroprotective; anorectic; tranquilliser; vulnerary;
 KW anorectic; cardiatic; nephrotropic; dermatological; antiHIV; antiviral;
 KW hyperglycaemia; obesity; lung disease; glomerulonephritis;
 KW interstitial nephritis; Turner's syndrome; Laron's syndrome;
 KW short stature; increased fat mass-to-lean ratio; immunological disorder;
 KW peripheral neuropathy; multiple sclerosis; muscular dystrophy;
 KW catabolic state; trauma; wounding; infection; HIV; skin disorder;
 KW human immunodeficiency virus; diabetes; heart dysfunction;
 KW kidney disorder; whole body growth disorder.

XX

OS Homo sapiens.

XX

PN WO200040612-A1.

XX

PD 13-JUL-2000.

XX

PF 05-JAN-2000; 2000WO-US000151.

XX
 PR 06-JAN-1999; 99US-0115010P.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Dubaquin Y, Lowman H;
 XX
 DR WPI; 2000-465955/40.
 XX
 PT Novel insulin-like growth factor (IGF) 1 mutants that selectively bind to
 PT IGF binding protein (IGFBP)-1 or IGFBP-3, used to improve the half-lives
 PT of IGF-I and insulin.
 XX
 PS Disclosure; Page 44; 48pp; English.
 XX
 CC The present invention describes an insulin-like growth factor (IGF)-1
 CC variant (I), where an amino acid at position 3, 4, 5, 7, 10, 14, 17, 23,
 CC 24, 25, 43, 49 or 63, optionally in combination with an amino acid at
 CC position 12 and/or 16 of the native human IGF-1 sequence, is replaced
 CC with an alanine, glycine, or a serine residue. The residue at position 7
 CC may be replaced by any amino acid. (I) can have antidiabetic, cardiant,
 CC neuroprotective, anorectic, tranquilliser, vulnerary, anorectic,
 CC nephrotropic, dermatological, antiHIV and antiviral activities. The IGF-1
 CC mutants are used in any methods where IGFs or insulin are used, e.g. in
 CC treating hyperglycaemia, obesity-related, neurological, cardiac, renal,
 CC immunological, and anabolic disorders. These disorders include lung
 CC diseases, glomerulonephritis, interstitial nephritis, Turner's syndrome,
 CC Laron's syndrome, short stature, increased fat mass-to-lean ratios,
 CC immunological disorders, peripheral neuropathy, multiple sclerosis,
 CC muscular dystrophy, catabolic states, trauma, wounding, infection, human
 CC immunodeficiency virus (HIV), wounds, skin disorders, diabetes, heart
 CC dysfunctions, kidney disorders, and whole body growth disorders. They can
 CC also be used for increasing serum and tissue levels of biological active
 CC IGF or insulin a mammal. The IGF-1 mutants improve the half-lives of IGF-
 CC 1 and insulin. The present sequence represents the native human
 CC proinsulin protein sequence, which is given in the exemplification of the
 CC present invention
 XX
 SQ Sequence 86 AA;

Query Match 100.0%; Score 463; DB 3; Length 86;
 Best Local Similarity 100.0%; Pred. No. 1.5e-43;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
 QY 61 SLQKRGIVEQCCTSICSLYQLENYCN 86
 ||||||||||||||||||||||||
 Db 61 SLQKRGIVEQCCTSICSLYQLENYCN 86

RESULT 5
 AAM48218
 ID AAM48218 standard; protein; 86 AA.
 XX

AC AAM48218;
 XX
 DT 18-MAR-2002 (first entry)
 XX
 DE Human proinsulin.
 XX
 KW Antirheumatic; antiarthritic; osteopathic; cartilage disorder;
 KW insulin-like growth factor; IGF; binding protein; IGFBP;
 KW rheumatoid arthritis; osteoarthritis; proinsulin; human.
 XX
 OS Homo sapiens.
 XX
 PN WO200187323-A2.
 XX
 PD 22-NOV-2001.
 XX
 PF 16-MAY-2001; 2001WO-US015904.
 XX
 PR 16-MAY-2000; 2000US-0204490P.
 PR 15-NOV-2000; 2000US-0248985P.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Dubaquié Y, Filvaroff EH, Lowman HB;
 XX
 DR WPI; 2002-082942/11.
 XX
 PT Treating cartilage disorders including cartilage damage by injury or
 PT degenerative cartilagenous disorders, by contacting cartilage with
 PT insulin-like growth factor analog with altered affinity for IGF-binding
 PT proteins.
 XX
 PS Disclosure; Fig 16; 136pp; English.
 XX
 CC The present invention relates to a method for treating cartilage
 CC disorders. The method comprises contacting cartilage with an active agent
 CC such as insulin-like growth factor (IGF-1) analog with a binding affinity
 CC preference for IGF binding protein-3 (IGFBP-3) over IGFBP-1, an IGF-1
 CC analog with a binding affinity preference for IGFBP-1 over IGFBP-3, or a
 CC IGFBP displacer peptide that prevents the interaction of IGF with an
 CC IGFBP and does not bind to human IGF receptor. The method is useful for
 CC treating cartilage disorders (CD), including degenerative CD, articular
 CC CD such as rheumatoid arthritis and osteoarthritis. The present sequence
 CC is human proinsulin, which was used to illustrate the invention
 XX
 SQ Sequence 86 AA;

Query Match 100.0%; Score 463; DB 5; Length 86;
 Best Local Similarity 100.0%; Pred. No. 1.5e-43;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60

Qy      61 SLQKRGIVEQCCTSICSLYQLENYCN 86
          ||||||||||||||||||||
  
```


RESULT 6

ADC64463

ID ADC64463 standard; protein; 86 AA.

XX

AC ADC64463;

XX

DT 18-DEC-2003 (first entry)

XX

DE Amino acid sequence for human proinsulin.

XX

KW Immunoassay; human C-peptide; HCP; immune complex; human; proinsulin.

XX

OS Homo sapiens.

XX

PN US2002160435-A1.

XX

PD 31-OCT-2002.

XX

PF 12-JUN-2001; 2001US-00878380.

XX

PR 12-JUN-2000; 2000JP-00174691.

XX

PA (KITA/) KITAJIMA S.

PA (KURA/) KURANO Y.

PA (NAKA/) NAKATSUBO K.

PA (NISH/) NISHIZONO I.

XX

PI Kitajima S, Kurano Y, Nakatsubo K, Nishizono I;

XX

DR WPI; 2003-765139/72.

XX

PT Measuring human C-peptide, by reacting sample C-peptide with two
PT different human C-peptide antibodies that recognize different epitopes on
PT peptide, to form immune complex, separating and quantifying immune
PT complex.

XX

PS Disclosure; SEQ ID NO 1; 20pp; English.

XX

CC The present invention relates to an immunoassay for measuring human C-
CC peptide (HCP). The method comprises reacting HCP in a sample with a first
CC anti-HCP antibody and a second anti-HCP antibody which is immobilised on
CC a support, to form an immune complex, and separating and quantifying the
CC immune complex, where the first and second antibody recognises the
CC epitope existing in the region from 1-110 and 1-16 amino acid residues,
CC respectively, from the N-terminal end of HCP. Also disclosed is a kit for
CC measuring human C-peptide. The method is useful for measuring human C-
CC peptides. The method provides high reproducibility, high detection
CC sensitivity, and low cross-reactivity to proinsulin. The present sequence
CC represents the amino acid sequence for human proinsulin.

XX

SQ Sequence 86 AA;

Query Match 100.0%; Score 463; DB 7; Length 86;

Best Local Similarity 100.0%; Pred. No. 1.5e-43;

Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60

Qy     61 SLQKRGIVEQCCTSICSLYQLENYCN 86
          ||||||||||||||||||||
Db     61 SLQKRGIVEQCCTSICSLYQLENYCN 86
```

RESULT 7

ADF16632

ID ADF16632 standard; protein; 86 AA.

XX

AC ADF16632;

XX

DT 12-FEB-2004 (first entry)

XX

DE Human albumin fusion protein-related protein SeqID1734.

XX

KW albumin fusion protein; albumin activity; human serum albumin;

KW serum osmotic pressure; shelf-life; stability; antidiabetic;

KW gene therapy; diabetes mellitus; human; gene; ds.

XX

OS Homo sapiens.

XX

PN WO2003060071-A2.

XX

PD 24-JUL-2003.

XX

PF 23-DEC-2002; 2002WO-US040891.

XX

PR 21-DEC-2001; 2001US-0341811P.

PR 24-JAN-2002; 2002US-0350358P.

PR 28-JAN-2002; 2002US-0351360P.

PR 26-FEB-2002; 2002US-0359370P.

PR 28-FEB-2002; 2002US-0360000P.

PR 27-MAR-2002; 2002US-0367500P.

PR 08-APR-2002; 2002US-0370227P.

PR 10-MAY-2002; 2002US-0378950P.

PR 24-MAY-2002; 2002US-0382617P.

PR 28-MAY-2002; 2002US-0383123P.

PR 05-JUN-2002; 2002US-0385708P.

PR 10-JUL-2002; 2002US-0394625P.

PR 24-JUL-2002; 2002US-0398008P.

PR 09-AUG-2002; 2002US-0402131P.

PR 13-AUG-2002; 2002US-0402708P.

PR 18-SEP-2002; 2002US-0411355P.

PR 18-SEP-2002; 2002US-0411426P.

PR 02-OCT-2002; 2002US-0414984P.

PR 11-OCT-2002; 2002US-0417611P.

PR 23-OCT-2002; 2002US-0420246P.

PR 05-NOV-2002; 2002US-0423623P.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

PA (DELZ) DELTA BIOTECHNOLOGY LTD.

PA (PRIN-) PRINCIPIA PHARM CORP.
 XX
 PI Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;
 XX
 DR WPI; 2003-598517/56.
 DR N-PSDB; ADF16306.
 XX
 PT New albumin fusion protein, useful for preparing a composition for
 PT treating diabetes mellitus.
 XX
 PS Example 4; SEQ ID NO 1734; 24pp; English.
 XX
 CC This invention relates to a novel albumin fusion protein having albumin
 CC or biological activity. Human serum albumin is responsible for a
 CC significant proportion of the osmotic pressure of serum and also
 CC functions as a carrier of endogenous and exogenous ligands. The fusion of
 CC albumin to a therapeutic protein may increase shelf-life and stability of
 CC the therapeutic protein. The albumin fusion protein of the invention may
 CC allow production of compositions with antidiabetic activity whilst the
 CC nucleotide sequence which encodes it may be useful for gene therapy. The
 CC albumin fusion protein is useful for preparing a composition for treating
 CC diabetes mellitus. The present sequence is that of a therapeutic protein
 CC which was fused with human albumin to create a novel albumin fusion
 CC protein of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification, but was obtained in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/publishedpct_sequences
 XX
 SQ Sequence 86 AA;

Query Match 100.0%; Score 463; DB 7; Length 86;
 Best Local Similarity 100.0%; Pred. No. 1.5e-43;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60

 Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 86
 ||||||||||||||||||
 Db 61 SLQKRGIVEQCCTSICSLYQLENYCN 86

RESULT 8

ADH21860

ID ADH21860 standard; protein; 86 AA.

XX

AC ADH21860;

XX

DT 11-MAR-2004 (first entry)

XX

DE Human long-acting insulin peptide, SEQ ID NO:657.

XX

KW Fusion protein; human serum albumin; HSA; therapeutic protein;
 KW shelf-life; in vitro biological activity; in vivo biological activity;
 KW metabolic disorder; endocrine disorder; diabetes; type 1; type 2;
 KW diabetes-related condition; hyperglycaemia; neural disorder; neuropathy;
 KW retinopathy; cardiovascular disorder; heart disease; renal disorder;

KW obesity; glucose level maintenance; weight loss; antidiabetic; cardiant;
 KW anorectic; ophthalmological; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO2003059934-A2.
 XX
 PD 24-JUL-2003.
 XX
 PF 23-DEC-2002; 2002WO-US040892.
 XX
 PR 21-DEC-2001; 2001US-0341811P.
 PR 24-JAN-2002; 2002US-0350358P.
 PR 26-FEB-2002; 2002US-0359370P.
 PR 28-FEB-2002; 2002US-0360000P.
 PR 27-MAR-2002; 2002US-0367500P.
 PR 08-APR-2002; 2002US-0370227P.
 PR 10-MAY-2002; 2002US-0378950P.
 PR 24-JUL-2002; 2002US-0398008P.
 PR 09-AUG-2002; 2002US-0402131P.
 PR 13-AUG-2002; 2002US-0402708P.
 PR 18-SEP-2002; 2002US-0411355P.
 PR 02-OCT-2002; 2002US-0414984P.
 PR 11-OCT-2002; 2002US-0417611P.
 PR 23-OCT-2002; 2002US-0420246P.
 PR 05-NOV-2002; 2002US-0423623P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Haseltine WA;
 XX
 DR WPI; 2003-598501/56.
 DR N-PSDB; ADH21708.
 XX
 PT New albumin fusion protein, useful for preparing a composition for
 PT treating diabetes mellitus.
 XX
 PS Disclosure; SEQ ID NO 657; 1086pp; English.
 XX
 CC The invention relates to fusion proteins comprising human serum albumin
 CC (ADH21530) and a therapeutic polypeptide such as a therapeutic protein,
 CC antibody or peptide or their variants or fragments. The therapeutic
 CC protein may be fused to the N-terminus, the C-terminus or both termini of
 CC albumin via a linker. The albumin component of the fusion proteins
 CC prolongs the shelf-life and the in vitro and vivo biological activity of
 CC the proteins compared with those of the corresponding therapeutic
 CC proteins on their own. The invention also relates to nucleic acids
 CC encoding albumin fusion proteins, vectors and host cells comprising an
 CC albumin fusion protein nucleic acid, compositions and kits comprising an
 CC albumin fusion protein, the method of extending the shelf-life of a
 CC therapeutic protein by fusion with albumin, and the treatment of disease
 CC using an albumin fusion protein. The albumin fusion proteins may be used
 CC in the treatment of metabolic/endocrine disorders, diabetes and diabetes-
 CC related conditions. Specifically the albumin fusion proteins may be used
 CC to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders
 CC (especially neuropathy), retinopathy, cardiovascular disorders
 CC (especially heart disease, renal disorders and obesity. The proteins may

CC also be used in a method of maintaining a basal glucose level in a
CC patient and in a method for losing weight. The present sequence is
CC related to the invention.

XX

SQ Sequence 86 AA;

Query Match 100.0%; Score 463; DB 7; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.5e-43;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
|||||

Db 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60

Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 86
|||||

Db 61 SLQKRGIVEQCCTSICSLYQLENYCN 86

RESULT 9

ADT93277

ID ADT93277 standard; protein; 86 AA.

XX

AC ADT93277;

XX

DT 16-DEC-2004 (first entry)

XX

DE Human native proinsulin protein.

XX

KW antidiabetic; nephrotropic; cardiovascular; hepatotropic; anabolic;

KW gene therapy; insulin-like growth factor-I; IGF-I; dysregulation;

KW GH/IGF axis; hyperglycemic disorder; renal disorder;

KW congestive heart failure; hepatic failure; poor nutrition;

KW wasting syndrome; catabolic state; IGF binding protein-1; IGFBP-1;

KW renal failure; proinsulin.

XX

OS Homo sapiens.

XX

PN AU2003236454-A1.

XX

PD 18-SEP-2003.

XX

PF 22-AUG-2003; 2003AU-00236454.

XX

PR 22-AUG-2003; 2003AU-00236454.

XX

PA (GETH) GENENTECH INC.

XX

PI Mortensen DL, Lowman HB, Fielder PJ, Dubaquié Y;

XX

DR WPI; 2004-662617/65.

XX

PT New insulin-like growth factor-I (IGF-I) variant, useful for treating

PT disorder associated with dysregulation of GH(growth hormone)/IGF axis

PT e.g. renal disorder.

XX

PS Disclosure; SEQ ID NO 2; 61pp; English.

PR 02-JAN-1982; 82US-00222010.
 PR 03-MAR-1982; 82US-00354287.
 XX
 PA (UYN-) STATE UNIV NEW YORK.
 XX
 PI Inouye M, Nakamura K;
 XX
 DR WPI; 1982-59775E/29.
 DR N-PSDB; AAN20041.
 XX
 PT Plasmid cloning vehicles - useful for transforming bacterial hosts to
 PT produce eukaryotic polypeptide(s).
 XX
 PS Disclosure; Fig 27; 114pp; English.
 XX
 CC The sequence comprises human proinsulin. (Updated on 25-MAR-2003 to
 CC correct PR field.)
 XX
 SQ Sequence 87 AA;

Query Match 100.0%; Score 463; DB 1; Length 87;
 Best Local Similarity 100.0%; Pred. No. 1.5e-43;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 61

 Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 86
 ||||||||||||||||||||||||||||
 Db 62 SLQKRGIVEQCCTSICSLYQLENYCN 87

RESULT 11

AAP40217

ID AAP40217 standard; protein; 87 AA.

XX

AC AAP40217;

XX

DT 25-MAR-2003 (revised)

DT 12-FEB-1992 (first entry)

XX

DE Sequence of the 32 N-terminal AAs of proinsulin.

XX

KW Hormone; cloning vector; phage resistant.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Region 2. .31

FT /label= B-chain

FT Region 32. .66

FT /label= C-chain

FT Region 67. .87

FT /label= A-chain

XX

PN GB2126237-A.


```

XX      Homo sapiens.
XX      Synthetic.
XX
FH      Key                Location/Qualifiers
FT      Region             2. .31
FT                        /label= A chain
FT      Region             32. .66
FT                        /label= B chain
FT      Region             67. .87
FT                        /label= A chain
XX
PN      EP154539-A.
XX
PD      11-SEP-1985.
XX
PF      04-MAR-1985;      85EP-00301469.
XX
PR      06-MAR-1984;      84US-00586592.
XX
PA      (ELIL ) LILLY & CO ELI.
XX
PI      Schoner R,  Schoner B;
XX
DR      WPI; 1985-224921/37.
DR      N-PSDB; AAN50152.
XX
PT      New recombinant DNA expression vector - with autonomous replication and
PT      on transcription generating polycistronic mrna.
XX
PS      Example; Fig 14; 118pp; English.
XX
CC      The inventors claim a process for preparing selectable and autonomously
CC      replicating recombinant DNA expression vectors which comprise 1) a
CC      transcriptional and translational activating sequence which is in the
CC      reading frame of a nucleotide sequence which codes for a peptide or
CC      polypeptide; 2) a translational stop signal; 3) a translational start
CC      signal which is in the reading frame of a nucleotide sequence that codes
CC      for a functional polypeptide; and 4) an additional translational stop
CC      signal. The peptide or polypeptide coding sequence codes for 2-20 AAs,
CC      esp. AAP50122-P50125. The functional polypeptide is esp. growth hormone,
CC      human insulin, interferon and human tissue plasminogen activator.
CC      (Updated on 16-AUG-2002 to add missing OS field.) (Updated on 25-MAR-2003
CC      to correct PA field.)
XX
SQ      Sequence 87 AA;

Query Match          100.0%;  Score 463;  DB 1;  Length 87;
Best Local Similarity 100.0%;  Pred. No. 1.5e-43;
Matches   86;  Conservative    0;  Mismatches    0;  Indels      0;  Gaps      0;

Qy          1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
             |||
Db          2 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 61
             |||

Qy          61 SLQKRGIVEQCCTSICSLYQLENYCN 86
             |||

```

RESULT 13

AAP50060

ID AAP50060 standard; protein; 87 AA.

XX

AC AAP50060;

XX

DT 25-MAR-2003 (revised)

DT 16-AUG-2002 (revised)

DT 11-NOV-1991 (first entry)

XX

DE Synthetic proinsulin.

XX

KW Proinsulin; vector; proteinaceous granule.

XX

OS Homo sapiens.

XX

| | | |
|----|-----|---------------------|
| FH | Key | Location/Qualifiers |
|----|-----|---------------------|

| | | |
|----|--------|--------|
| FT | Region | 1. .30 |
|----|--------|--------|

| | | |
|----|--|------------------|
| FT | | /label= B chain. |
|----|--|------------------|

| | | |
|----|--------|---------|
| FT | Region | 31. .65 |
|----|--------|---------|

| | | |
|----|--|------------------|
| FT | | /label= C chain. |
|----|--|------------------|

| | | |
|----|--------|---------|
| FT | Region | 66. .86 |
|----|--------|---------|

| | | |
|----|--|------------------|
| FT | | /label= A chain. |
|----|--|------------------|

XX

PN EP159123-A.

XX

PD 23-OCT-1985.

XX

PF 04-MAR-1985; 85EP-00301468.

XX

PR 06-MAR-1984; 84US-00586582.

PR 26-JUL-1984; 84US-00634920.

PR 31-JAN-1985; 85US-00697090.

XX

PA (ELIL) LILLY & CO.ELI.

XX

PI Hsiung HM, Schoner RG, Schoner BE;

XX

DR WPI; 1985-265090/43.

DR N-PSDB; AAN50082.

XX

PT New selectable and autonomously replicating DNA expression vector -
PT useful in producing proteinaceous granules in cell transformants, esp.
PT for prodn. of bovine growth hormone derivs.

XX

PS Disclosure; Fig 14; 115pp; English.

XX

CC The synthetic proinsulin gene is expressed in a new selectable and
CC autonomously replicating recombinant DNA expression vector comprising a
CC runaway replicon and a transcriptional and translational activating
CC sequence in the reading frame of the proinsulin coding sequence, the
CC sequence contg. a translational stop signal. Host cells contg. the
CC vector, which is esp. plasmid pCZ103, are cultured, and proinsulin is
CC produced as a highly homogeneous species of proteinaceous granule. The

CC granule can be readily isolated from cell lysates and is stable on
CC washing with urea or detergent solns. at low concns. The granule contains
CC at least 50% of proinsulin and all isolation operations are simplified.
CC (Updated on 16-AUG-2002 to add missing OS field.) (Updated on 25-MAR-2003
CC to correct PA field.)

XX

SQ Sequence 87 AA;

Query Match 100.0%; Score 463; DB 1; Length 87;
Best Local Similarity 100.0%; Pred. No. 1.5e-43;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 2 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 61

Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 86
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 62 SLQKRGIVEQCCTSICSLYQLENYCN 87

RESULT 14

AAP61090

ID AAP61090 standard; protein; 87 AA.

XX

AC AAP61090;

XX

DT 28-FEB-1992 (first entry)

XX

DE Sequence encoded by the structural gene for human proinsulin.

XX

KW Recombinant plasmid; E.coli expression vector; secretion vector.

XX

OS Homo sapiens.

XX

PN US4624926-A.

XX

PD 25-NOV-1986.

XX

PF 03-MAR-1982; 82US-00354287.

XX

PR 02-JAN-1981; 81US-00222010.

PR 23-JUL-1981; 81US-00286070.

XX

PA (UYNY-) UNIV OF NEW YORK.

XX

PI Inouye M, Nakamura K;

XX

DR WPI; 1986-331802/50.

DR N-PSDB; AAN60872.

XX

PT New recombinant plasmid(s) - contg. DNA sequences encoding exogenous
PT polypeptide and outer membrane protein of E coli.

XX

PS Example; Fig 27; 44pp; English.

XX

CC The inventors claim new recombinant plasmids contg. a DNA sequence

CC encoding a polypeptide, which is foreign to E.coli, in reading phase with
CC a DNA SQ, coding for at least one functional fragment derived from an
CC outer membrane lipoprotein gene of E.coli. The foreign gene may be for
CC human insulin. The lipoprotein gene functional fragment may be the
CC promoter, the 5'-UTR, the 3'-UTR or the transcription termination signal
CC provided that it includes at least the promoter

XX

SQ Sequence 87 AA;

Query Match 100.0%; Score 463; DB 1; Length 87;

Best Local Similarity 100.0%; Pred. No. 1.5e-43;

Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
|||||

Db 2 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 61

Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 86
|||||

Db 62 SLQKRGIVEQCCTSICSLYQLENYCN 87

RESULT 15

AAR32367

ID AAR32367 standard; protein; 87 AA.

XX

AC AAR32367;

XX

DT 25-MAR-2003 (revised)

DT 18-JUN-1993 (first entry)

XX

DE Proinsulin protein sequence.

XX

KW Human; proinsulin; vector; pUC19; pPINS; CAT; pUC-CAT-proinsulin;

KW insulin analogue; type I; type II; diabetes.

XX

OS Synthetic.

XX

PN WO9303174-A1.

XX

PD 18-FEB-1993.

XX

PF 31-JUL-1992; 92WO-US006451.

XX

PR 08-AUG-1991; 91US-00741938.

PR 30-JUL-1992; 92US-00918953.

XX

PA (SCIO-) SCIOS INC.

PA (PFIZ) PFIZER INC.

XX

PI Andy RJ, Larson ER;

XX

DR WPI; 1993-076530/09.

DR N-PSDB; AAQ37003.

XX

PT New hepato selective and peripheral selective human insulin analogues -

PT and their corresp. DNA, for treatment of type I and type II diabetes.

XX
 PS Disclosure; Fig 2b; 58pp; English.
 XX
 CC This sequence represents human proinsulin and was decoded from the
 CC sequences given in AAQ36996-7001. The cDNA fragment coding for proinsulin
 CC was inserted into plasmid vector pUC19 and digested with KpnI and
 CC HindIII. This resulted in the formation of the vector pPINS. A fragment
 CC encoding amino acids 1-73 of CAT (see AAQ37002) was inserted into pPINS
 CC to give a plasmid which contained DNA sequences which coded for amino
 CC acids 1-73 of CAT, an 8 amino acid linker sequence and human proinsulin.
 CC This plasmid, pUC-CAT-proinsulin, could be used in the formation of
 CC insulin analogues which may be used in the treatment of types I and II
 CC diabetes. (Updated on 25-MAR-2003 to correct PN field.)

XX
 SQ Sequence 87 AA;

Query Match 100.0%; Score 463; DB 2; Length 87;
 Best Local Similarity 100.0%; Pred. No. 1.5e-43;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|----|--|----|
| Qy | 1 | FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG | 60 |
| | | | |
| Db | 2 | FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG | 61 |
| Qy | 61 | SLQKRGIVEQCCTSICSLYQLENYCN | 86 |
| | | | |
| Db | 62 | SLQKRGIVEQCCTSICSLYQLENYCN | 87 |

Search completed: February 11, 2005, 18:14:51
 Job time : 92.0148 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2005, 18:04:56 ; Search time 22.69 Seconds
(without alignments)
282.936 Million cell updates/sec

Title: US-10-054-873-4
Perfect score: 463
Sequence: 1 FVNQHLCGSHLVEALYLVCG.....IVEQCCTSICSLYQLENYCN 86

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query | | DB | ID | Description |
|---------------|-------|-------|--------|----|-------------------|--------------------|
| | | Match | Length | | | |
| 1 | 463 | 100.0 | 86 | 4 | US-09-477-924-2 | Sequence 2, Appli |
| 2 | 463 | 100.0 | 86 | 4 | US-09-723-981-2 | Sequence 2, Appli |
| 3 | 463 | 100.0 | 86 | 4 | US-09-723-896-2 | Sequence 2, Appli |
| 4 | 463 | 100.0 | 86 | 4 | US-09-878-380-1 | Sequence 1, Appli |
| 5 | 463 | 100.0 | 96 | 2 | US-09-134-836-4 | Sequence 4, Appli |
| 6 | 463 | 100.0 | 96 | 3 | US-09-386-303A-4 | Sequence 4, Appli |
| 7 | 463 | 100.0 | 96 | 4 | US-09-947-563-4 | Sequence 4, Appli |
| 8 | 463 | 100.0 | 97 | 1 | US-08-160-376A-4 | Sequence 4, Appli |
| 9 | 463 | 100.0 | 110 | 3 | US-08-950-720A-11 | Sequence 11, Appli |
| 10 | 463 | 100.0 | 110 | 3 | US-08-589-028-2 | Sequence 2, Appli |
| 11 | 463 | 100.0 | 110 | 3 | US-08-784-582-2 | Sequence 2, Appli |

| | | | | | | |
|----|-------|-------|-----|---|-------------------|-------------------|
| 12 | 463 | 100.0 | 110 | 3 | US-08-785-271-2 | Sequence 2, Appli |
| 13 | 463 | 100.0 | 110 | 4 | US-08-472-701-2 | Sequence 2, Appli |
| 14 | 463 | 100.0 | 110 | 4 | US-09-185-852-2 | Sequence 2, Appli |
| 15 | 463 | 100.0 | 110 | 4 | US-09-815-229-3 | Sequence 3, Appli |
| 16 | 463 | 100.0 | 110 | 4 | US-09-617-389B-20 | Sequence 20, Appl |
| 17 | 463 | 100.0 | 110 | 4 | US-09-323-738-2 | Sequence 2, Appli |
| 18 | 463 | 100.0 | 110 | 4 | US-09-015-399-7 | Sequence 7, Appli |
| 19 | 463 | 100.0 | 110 | 5 | PCT-US95-08596-2 | Sequence 2, Appli |
| 20 | 463 | 100.0 | 117 | 4 | US-09-280-030-63 | Sequence 63, Appl |
| 21 | 463 | 100.0 | 130 | 4 | US-09-280-030-62 | Sequence 62, Appl |
| 22 | 463 | 100.0 | 151 | 2 | US-08-508-664-15 | Sequence 15, Appl |
| 23 | 463 | 100.0 | 161 | 2 | US-08-508-664-16 | Sequence 16, Appl |
| 24 | 463 | 100.0 | 167 | 1 | US-07-918-953-8 | Sequence 8, Appli |
| 25 | 463 | 100.0 | 167 | 1 | US-08-081-661-8 | Sequence 8, Appli |
| 26 | 457 | 98.7 | 96 | 2 | US-09-134-836-5 | Sequence 5, Appli |
| 27 | 457 | 98.7 | 96 | 3 | US-09-386-303A-5 | Sequence 5, Appli |
| 28 | 457 | 98.7 | 96 | 4 | US-09-947-563-5 | Sequence 5, Appli |
| 29 | 457 | 98.7 | 97 | 1 | US-08-389-487-7 | Sequence 7, Appli |
| 30 | 456 | 98.5 | 90 | 1 | US-08-030-731A-43 | Sequence 43, Appl |
| 31 | 456 | 98.5 | 98 | 4 | US-09-701-968-7 | Sequence 7, Appli |
| 32 | 456 | 98.5 | 99 | 4 | US-09-701-968-8 | Sequence 8, Appli |
| 33 | 456 | 98.5 | 100 | 4 | US-09-701-968-9 | Sequence 9, Appli |
| 34 | 449 | 97.0 | 110 | 4 | US-09-574-443-1 | Sequence 1, Appli |
| 35 | 446 | 96.3 | 97 | 3 | US-09-099-307-6 | Sequence 6, Appli |
| 36 | 444 | 95.9 | 97 | 3 | US-09-099-307-8 | Sequence 8, Appli |
| 37 | 443 | 95.7 | 110 | 3 | US-08-589-028-4 | Sequence 4, Appli |
| 38 | 443 | 95.7 | 110 | 3 | US-08-784-582-4 | Sequence 4, Appli |
| 39 | 443 | 95.7 | 110 | 3 | US-08-785-271-4 | Sequence 4, Appli |
| 40 | 440 | 95.0 | 97 | 3 | US-09-099-307-7 | Sequence 7, Appli |
| 41 | 435 | 94.0 | 97 | 3 | US-09-099-307-11 | Sequence 11, Appl |
| 42 | 398 | 86.0 | 91 | 4 | US-09-676-787-7 | Sequence 7, Appli |
| 43 | 300 | 64.8 | 56 | 4 | US-09-815-229-10 | Sequence 10, Appl |
| 44 | 292.5 | 63.2 | 67 | 3 | US-08-981-988A-1 | Sequence 1, Appli |
| 45 | 290.5 | 62.7 | 83 | 3 | US-08-981-988A-3 | Sequence 3, Appli |

ALIGNMENTS

RESULT 1

US-09-477-924-2

```

; Sequence 2, Application US/09477924
; Patent No. 6403764
; GENERAL INFORMATION:
; APPLICANT: Dubaquie, Yves
; APPLICANT: Lowman, Henry
; TITLE OF INVENTION: PROTEIN VARIANTS
; FILE REFERENCE: P1712R1-1
; CURRENT APPLICATION NUMBER: US/09/477,924
; CURRENT FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 2
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-477-924-2

```

Query Match 100.0%; Score 463; DB 4; Length 86;
Best Local Similarity 100.0%; Pred. No. 2.4e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60

Qy      61 SLQKRGIVEQCCTSICSLYQLENYCN 86
          ||||||||||||||||||
Db      61 SLQKRGIVEQCCTSICSLYQLENYCN 86
```

RESULT 2

US-09-723-981-2
; Sequence 2, Application US/09723981
; Patent No. 6506874
; GENERAL INFORMATION:
; APPLICANT: Dubaquier, Yves
; APPLICANT: Lowman, Henry
; TITLE OF INVENTION: PROTEIN VARIANTS
; FILE REFERENCE: P1712R1
; CURRENT APPLICATION NUMBER: US/09/723,981
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/477,923
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 2
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-723-981-2

Query Match 100.0%; Score 463; DB 4; Length 86;
Best Local Similarity 100.0%; Pred. No. 2.4e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60

Qy      61 SLQKRGIVEQCCTSICSLYQLENYCN 86
          ||||||||||||||||||
Db      61 SLQKRGIVEQCCTSICSLYQLENYCN 86
```

RESULT 3

US-09-723-896-2
; Sequence 2, Application US/09723896
; Patent No. 6509443
; GENERAL INFORMATION:
; APPLICANT: Dubaquier, Yves
; APPLICANT: Lowman, Henry
; TITLE OF INVENTION: PROTEIN VARIANTS
; FILE REFERENCE: P1712R1
; CURRENT APPLICATION NUMBER: US/09/723,896
; CURRENT FILING DATE: 2000-11-28

Db

61 SLQKRGIVEQCCTSIQSLYQLENYCN 86

RESULT 5

US-09-134-836-4

; Sequence 4, Application US/09134836

; Patent No. 5986048

; GENERAL INFORMATION:

; APPLICANT: Rubroder, Franz-Josef

; APPLICANT: Keller, Reinhold

; TITLE OF INVENTION: Improved process for obtaining

; TITLE OF INVENTION: insulin precursors having correctly bonded cystine
bridges

; NUMBER OF SEQUENCES: 7

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Finnegan, Henderson, Farrahaw, Garrett &

; ADDRESSEE: Dunner

; STREET: 1300 I Street, N.W.

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20005-3315

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/134,836

; FILING DATE:

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Leslie McDonell

; REGISTRATION NUMBER: 34,872

; REFERENCE/DOCKET NUMBER: 02481.1600-00000

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 408-4000

; TELEFAX: (202) 408-4400

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 96 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; ORIGINAL SOURCE:

; ORGANISM: Escherichia coli

; FEATURE:

; NAME/KEY: Protein

; LOCATION: 1..96

US-09-134-836-4

Query Match 100.0%; Score 463; DB 2; Length 96;

Best Local Similarity 100.0%; Pred. No. 2.8e-47;

Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy

1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60

```

Db      11 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 70
Qy      61 SLQKRGIVEQCCTSICSLYQLENYCN 86
Db      71 SLQKRGIVEQCCTSICSLYQLENYCN 96

```

RESULT 6

US-09-386-303A-4

; Sequence 4, Application US/09386303A

; Patent No. 6380355

; GENERAL INFORMATION:

; APPLICANT: Rubroder, Franz-Josef

; Keller, Reinhold

; TITLE OF INVENTION: Improved process for obtaining

; insulin precursors having correctly bonded cystine

bridges

; NUMBER OF SEQUENCES: 7

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Finnegan, Henderson, Farrabow, Garrett &

; Dunner

; STREET: 1300 I Street, N.W.

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20005-3315

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/386,303A

; FILING DATE: 31-Aug-1999

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/134,836

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Leslie McDonell

; REGISTRATION NUMBER: 34,872

; REFERENCE/DOCKET NUMBER: 02481.1600-00000

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 408-4000

; TELEFAX: (202) 408-4400

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 96 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; ORIGINAL SOURCE:

; ORGANISM: Escherichia coli

; FEATURE:

; NAME/KEY: Protein

; LOCATION: 1..96
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-386-303A-4

Query Match 100.0%; Score 463; DB 3; Length 96;
Best Local Similarity 100.0%; Pred. No. 2.8e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
|
Db 11 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 70

Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 86
|
Db 71 SLQKRGIVEQCCTSICSLYQLENYCN 96

RESULT 7

US-09-947-563-4

; Sequence 4, Application US/09947563
; Patent No. 6727346

; GENERAL INFORMATION:

; APPLICANT: Rubroder, Franz-Josef
; Keller, Reinhold

; TITLE OF INVENTION: Improved process for obtaining
; insulin precursors having correctly bonded cystine
bridges

; NUMBER OF SEQUENCES: 7

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Finnegan, Henderson, Farrahaw, Garrett &
; Dunner

; STREET: 1300 I Street, N.W.

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20005-3315

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/947,563

; FILING DATE: 07-Sep-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/134,836

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Leslie McDonell

; REGISTRATION NUMBER: 34,872

; REFERENCE/DOCKET NUMBER: 02481.1600-00000

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 408-4000

; TELEFAX: (202) 408-4400

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

```

;          LENGTH: 96 amino acids
;          TYPE: amino acid
;          STRANDEDNESS: single
;          TOPOLOGY: linear
;          MOLECULE TYPE: protein
;          ORIGINAL SOURCE:
;             ORGANISM: Escherichia coli
;          FEATURE:
;             NAME/KEY: Protein
;             LOCATION: 1..96
;          SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-947-563-4

```

```

Query Match          100.0%; Score 463; DB 4; Length 96;
Best Local Similarity 100.0%; Pred. No. 2.8e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy          1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
            |
Db          11 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 70

Qy          61 SLQKRGIVEQCCTSICSLYQLENYCN 86
            |
Db          71 SLQKRGIVEQCCTSICSLYQLENYCN 96

```

```

RESULT 8
US-08-160-376A-4
; Sequence 4, Application US/08160376A
; Patent No. 5473049
; GENERAL INFORMATION:
;   APPLICANT: Obermeier, Ranier
;   APPLICANT: Gerl, Martin
;   APPLICANT: Ludwig, Jurgen
;   APPLICANT: Sabel, Walter
;   TITLE OF INVENTION: Process For Obtaining Proinsulin
;   TITLE OF INVENTION: Possessing Correctly Linked
;   TITLE OF INVENTION: Cystine Bridges
;   NUMBER OF SEQUENCES: 7
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Kenneth A. Genoni, Esq.
;   STREET: Rt. 202-206 No. 5473049th/P.O. Box 2500
;   CITY: Somerville
;   STATE: New Jersey
;   COUNTRY: U.S.A.
;   ZIP: 08876-1258
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
;   COMPUTER: IBM 386
;   OPERATING SYSTEM: WINDOWS 3.1
;   SOFTWARE: WORDPERFECT 5.1
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/160,376A
;   FILING DATE: December 1, 1993
;   CLASSIFICATION: 530
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: GE P 4240420.7

```

```

; FILING DATE: December 2, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Barbara V. Maurer, Esq.
; REGISTRATION NUMBER: 31,287
; REFERENCE/DOCKET NUMBER: HOE 92/F 384
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 231-4079
; TELEFAX: (908) 231-2255
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 97 Amino Acids
; TYPE: Amino Acid (AA)
; TOPOLOGY: not relevant
US-08-160-376A-4

```

```

Query Match          100.0%; Score 463; DB 1; Length 97;
Best Local Similarity 100.0%; Pred. No. 2.8e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 71

Qy      61 SLQKRGIVEQCCTSICSLYQLENYCN 86
        ||||||||||||||||||||||||
Db      72 SLQKRGIVEQCCTSICSLYQLENYCN 97

```

RESULT 9

US-08-950-720A-11

; Sequence 11, Application US/08950720A

; Patent No. 6046028

; GENERAL INFORMATION:

; APPLICANT: Conklin, Darrell C.

; APPLICANT: Lofton-Day, Catherine E.

; APPLICANT: Lok, Si

; APPLICANT: Jaspers, Stephen R.

; TITLE OF INVENTION: INSULIN HOMOLOG

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ZymoGenetics, Inc.

; STREET: 1201 Eastlake Avenue East

; CITY: Seattle

; STATE: WA

; COUNTRY: USA

; ZIP: 98102

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/950,720A

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

```

; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sawislak, Deborah A
; REGISTRATION NUMBER: 37,438
; REFERENCE/DOCKET NUMBER: 96-09
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6672
; TELEFAX: 206-442-6678
; TELEX:
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 110 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6046028e
US-08-950-720A-11

```

```

Query Match          100.0%; Score 463; DB 3; Length 110;
Best Local Similarity 100.0%; Pred. No. 3.3e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 84

Qy      61 SLQKRGIVEQCCTSICSLYQLENYCN 86
          ||||||||||||||||||
Db      85 SLQKRGIVEQCCTSICSLYQLENYCN 110

```

RESULT 10

US-08-589-028-2

; Sequence 2, Application US/08589028

; Patent No. 6087129

; GENERAL INFORMATION:

; APPLICANT: Newgard, Christopher B.

; APPLICANT: Halban, Philippe

; APPLICANT: No. 6087129mington, Karl D.

; APPLICANT: Clark, Samuel A.

; APPLICANT: Thigpen, Anice E.

; APPLICANT: Quaade, Christian

; APPLICANT: Kruse, Fred

; TITLE OF INVENTION: Recombinant Expression of Proteins From

; TITLE OF INVENTION: Secretory Cell Lines

; NUMBER OF SEQUENCES: 50

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Arnold, White & Durkee

; STREET: P. O. Box 4433

; CITY: Houston

; STATE: TX

; COUNTRY: USA

; ZIP: 77210-4433

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

```

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/589,028
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 47,642
; REFERENCE/DOCKET NUMBER: UTSD:426\HYL
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 110 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-589-028-2

```

```

Query Match          100.0%; Score 463; DB 3; Length 110;
Best Local Similarity 100.0%; Pred. No. 3.3e-47;
Matches   86; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```

```

Qy      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 84

Qy      61 SLQKRGIVEQCCTSICSLYQLENYCN 86
        ||||||||||||||||||||
Db      85 SLQKRGIVEQCCTSICSLYQLENYCN 110

```

RESULT 11

US-08-784-582-2

```

; Sequence 2, Application US/08784582
; Patent No. 6110707
; GENERAL INFORMATION:
; APPLICANT: Newgard, Christopher B.
; APPLICANT: Halban, Philippe A.
; APPLICANT: No. 6110707mington, Karl D.
; APPLICANT: Clark, Samuel A.
; APPLICANT: Thigpen, Anice E.
; APPLICANT: Quaade, Christian
; APPLICANT: Kruse, Fred
; APPLICANT: McGarry, Dennis
; TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM
; TITLE OF INVENTION: SECRETORY CELL LINES
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:

```



```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/784,582
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/028,427
; FILING DATE: 15-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/589,028
; FILING DATE: 19-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: UTSD:514
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 110 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-784-582-2

```

```

Query Match          100.0%; Score 463; DB 3; Length 110;
Best Local Similarity 100.0%; Pred. No. 3.3e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 84

Qy      61 SLQKRGIVEQCCTSICSLYQLENYCN 86
        ||||||||||||||||||||||||
Db      85 SLQKRGIVEQCCTSICSLYQLENYCN 110

```

RESULT 12

US-08-785-271-2

; Sequence 2, Application US/08785271

; Patent No. 6194176

; GENERAL INFORMATION:

; APPLICANT: Newgard, Christopher B.

; APPLICANT: Halban, Philippe A.

; APPLICANT: No. 6194176mington, Karl D.

; APPLICANT: Clark, Samuel A.

; APPLICANT: Thigpen, Anice E.

; APPLICANT: Quaade, Christian

; APPLICANT: Kruse, Fred

; TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM

; TITLE OF INVENTION: SECRETORY CELL LINES

; NUMBER OF SEQUENCES: 56

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/785,271
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/589,028
; FILING DATE: 19-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: UTSD:513
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 110 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-785-271-2

```

```

Query Match          100.0%; Score 463; DB 3; Length 110;
Best Local Similarity 100.0%; Pred. No. 3.3e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 84

Qy      61 SLQKRGIVEQCCTSICSLYQLENYCN 86
        ||||||||||||||||||||
Db      85 SLQKRGIVEQCCTSICSLYQLENYCN 110

```

RESULT 13

US-08-472-701-2

```

; Sequence 2, Application US/08472701
; Patent No. 6509165
; GENERAL INFORMATION:
; APPLICANT: Griffin, Ann C.
; APPLICANT: Hickey, William F.
; TITLE OF INVENTION: Detection and Treatment Methods for
; TITLE OF INVENTION: Type I Diabetes
; NUMBER OF SEQUENCES: 23

```

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,701
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/272,220
; FILING DATE: 08-JULY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: DeConti, Giulio A., Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: DCI-092DV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 110 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-472-701-2

```

```

Query Match          100.0%;  Score 463;  DB 4;  Length 110;
Best Local Similarity 100.0%;  Pred. No. 3.3e-47;
Matches   86;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

```

```

Qy      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 84

Qy      61 SLQKRGIVEQCCTSICSLYQLENYCN 86
        ||||||||||||||||||||
Db      85 SLQKRGIVEQCCTSICSLYQLENYCN 110

```

```

RESULT 14
US-09-185-852-2
; Sequence 2, Application US/09185852
; Patent No. 6537806
; GENERAL INFORMATION:
; APPLICANT: Osborne, William R.A.
; APPLICANT: Ramesh, Nagarajan
; TITLE OF INVENTION: Compositions and Methods for Treating Diabetes
; FILE REFERENCE: P-UW 3264

```

```
; CURRENT APPLICATION NUMBER: US/09/185,852
; CURRENT FILING DATE: 1998-11-04
; EARLIER APPLICATION NUMBER: 60/087,660
; EARLIER FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 110
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-185-852-2
```

Query Match 100.0%; Score 463; DB 4; Length 110;
Best Local Similarity 100.0%; Pred. No. 3.3e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
      |||
Db      25 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 84
      |||
Qy      61 SLQKRGIVEQCCTSICSLYQLENYCN 86
      |||
Db      85 SLQKRGIVEQCCTSICSLYQLENYCN 110
```

RESULT 15

US-09-815-229-3

; Sequence 3, Application US/09815229

; Patent No. 6689747

; GENERAL INFORMATION:

; APPLICANT: Filvaroff, Ellen H.

; APPLICANT: Okumu, Franklin W.

; TITLE OF INVENTION: USE OF INSULIN FOR THE TREATMENT OF CARTILAGENOUS DISORDERS

; FILE REFERENCE: P1786R1US

; CURRENT APPLICATION NUMBER: US/09/815,229

; CURRENT FILING DATE: 2001-03-22

; PRIOR APPLICATION NUMBER: US 60/192,103

; PRIOR FILING DATE: 2000-03-24

; NUMBER OF SEQ ID NOS: 17

; SEQ ID NO 3

; LENGTH: 110

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-815-229-3

Query Match 100.0%; Score 463; DB 4; Length 110;
Best Local Similarity 100.0%; Pred. No. 3.3e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
      |||
Db      25 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 84
      |||
Qy      61 SLQKRGIVEQCCTSICSLYQLENYCN 86
      |||
Db      85 SLQKRGIVEQCCTSICSLYQLENYCN 110

```

Search completed: February 11, 2005, 18:27:05
Job time : 24.69 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2005, 17:42:33 ; Search time 16.3432 Seconds
(without alignments)
506.306 Million cell updates/sec

Title: US-10-054-873-4
Perfect score: 463
Sequence: 1 FVNQHLCGSHLVEALYLVCG.....IVEQCCTSICSLYQLENYCN 86

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*

- 1: pir1:*
- 2: pir2:*
- 3: pir3:*
- 4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % | | DB | ID | Description |
|------------|-------|-------------|--------|----|--------|-------------------|
| | | Query Match | Length | | | |
| 1 | 463 | 100.0 | 110 | 1 | IPHU | insulin precursor |
| 2 | 463 | 100.0 | 110 | 2 | A42179 | insulin precursor |
| 3 | 456 | 98.5 | 110 | 2 | B42179 | insulin precursor |
| 4 | 456 | 98.5 | 110 | 2 | JQ0178 | insulin precursor |
| 5 | 424 | 91.6 | 110 | 1 | INRB | insulin precursor |
| 6 | 417 | 90.1 | 110 | 1 | IPDG | insulin precursor |
| 7 | 394 | 85.1 | 86 | 1 | IPHO | insulin precursor |
| 8 | 394 | 85.1 | 110 | 1 | INMS2 | insulin 2 precurs |
| 9 | 394 | 85.1 | 110 | 1 | IPRT2 | insulin 2 precurs |
| 10 | 392 | 84.7 | 108 | 2 | A39883 | insulin precursor |
| 11 | 392 | 84.7 | 110 | 2 | I48166 | insulin precursor |
| 12 | 385 | 83.2 | 110 | 1 | IPRT1 | insulin 1 precurs |
| 13 | 383 | 82.7 | 84 | 1 | IPPG | insulin precursor |

| | | | | | | |
|----|-------|------|-----|---|--------|--------------------|
| 14 | 366.5 | 79.2 | 105 | 1 | IPBO | insulin precursor |
| 15 | 366 | 79.0 | 108 | 1 | INMS1 | insulin 1 precurs |
| 16 | 334.5 | 72.2 | 108 | 2 | S09278 | insulin precursor |
| 17 | 320.5 | 69.2 | 77 | 1 | INSH | insulin precursor |
| 18 | 314 | 67.8 | 110 | 1 | IPGP | insulin precursor |
| 19 | 277.5 | 59.9 | 109 | 1 | IPRTDU | insulin precursor |
| 20 | 276.5 | 59.7 | 103 | 2 | I51221 | insulin precursor |
| 21 | 265.5 | 57.3 | 106 | 1 | IPXL2 | insulin II precurs |
| 22 | 265.5 | 57.3 | 107 | 1 | IPCH | insulin precursor |
| 23 | 262.5 | 56.7 | 106 | 1 | IPXL1 | insulin I precurs |
| 24 | 256.5 | 55.4 | 51 | 1 | INEL | insulin - elephant |
| 25 | 256.5 | 55.4 | 51 | 1 | INWHF | insulin - finback |
| 26 | 256.5 | 55.4 | 51 | 1 | INWHP | insulin - sperm wh |
| 27 | 256.5 | 55.4 | 81 | 1 | IPDK | insulin precursor |
| 28 | 256 | 55.3 | 96 | 2 | PC7082 | epidermal growth f |
| 29 | 254.5 | 55.0 | 51 | 1 | INHY | insulin - hamster |
| 30 | 251.5 | 54.3 | 51 | 1 | INMSSP | insulin - Egyptian |
| 31 | 250.5 | 54.1 | 51 | 2 | A59151 | insulin precursor |
| 32 | 246.5 | 53.2 | 51 | 1 | INCMA | insulin - Arabian |
| 33 | 246.5 | 53.2 | 51 | 1 | INGT | insulin - goat |
| 34 | 246.5 | 53.2 | 51 | 1 | INWH1S | insulin - sei whal |
| 35 | 245.5 | 53.0 | 51 | 1 | INCT | insulin - cat |
| 36 | 244.5 | 52.8 | 51 | 1 | INMKSQ | insulin - common s |
| 37 | 239.5 | 51.7 | 51 | 2 | JQ0362 | insulin - North Am |
| 38 | 234.5 | 50.6 | 51 | 1 | INCB | insulin - Chinchil |
| 39 | 231.5 | 50.0 | 51 | 1 | INGS | insulin - goose |
| 40 | 227.5 | 49.1 | 51 | 1 | INOS | insulin - ostrich |
| 41 | 227.5 | 49.1 | 51 | 1 | INTK | insulin - turkey (|
| 42 | 227.5 | 49.1 | 51 | 1 | A61129 | insulin - black-be |
| 43 | 227.5 | 49.1 | 51 | 1 | INPQ | insulin - crested |
| 44 | 227.5 | 49.1 | 51 | 2 | A60414 | insulin - slider t |
| 45 | 225 | 48.6 | 52 | 2 | S44470 | insulin I2 - North |

ALIGNMENTS

RESULT 1

IPHU

insulin precursor [validated] - human

N;Alternate names: preproinsulin

C;Species: Homo sapiens (man)

C;Date: 23-Oct-1981 #sequence_revision 23-Oct-1981 #text_change 09-Jul-2004

C;Accession: A93222; A94253; A93216; A94251; A93144; A92075; A91186; I58114; A01579; S58661

R;Bell, G.I.; Pictet, R.L.; Rutter, W.J.; Cordell, B.; Tischer, E.; Goodman, H.M.

Nature 284, 26-32, 1980

A;Title: Sequence of the human insulin gene.

A;Reference number: A93222; MUID:80120725; PMID:6243748

A;Accession: A93222

A;Molecule type: DNA

A;Residues: 1-110 <BEL>

A;Cross-references: UNIPROT:P01308; GB:J00265; NID:g186429; PIDN:AAA59172.1; PID:g386828

R;Ullrich, A.; Dull, T.J.; Gray, A.; Brosius, J.; Sures, I.

Science 209, 612-615, 1980

A;Title: Genetic variation in the human insulin gene.
 A;Reference number: A94253; MUID:80236313; PMID:6248962
 A;Accession: A94253
 A;Molecule type: DNA
 A;Residues: 1-110 <ULL>
 A;Cross-references: GB:J00265; NID:g186429; PIDN:AAA59172.1; PID:g386828
 R;Bell, G.I.; Swain, W.F.; Pictet, R.; Cordell, B.; Goodman, H.M.; Rutter, W.J.
 Nature 282, 525-527, 1979
 A;Title: Nucleotide sequence of a cDNA clone encoding human preproinsulin.
 A;Reference number: A93216; MUID:80054779; PMID:503234
 A;Accession: A93216
 A;Molecule type: mRNA
 A;Residues: 1-110 <BEL2>
 A;Cross-references: GB:J00265; NID:g186429; PIDN:AAA59172.1; PID:g386828
 R;Sures, I.; Goeddel, D.V.; Gray, A.; Ullrich, A.
 Science 208, 57-59, 1980
 A;Title: Nucleotide sequence of human preproinsulin complementary DNA.
 A;Reference number: A94251; MUID:80147417; PMID:6927840
 A;Accession: A94251
 A;Molecule type: mRNA
 A;Residues: 1-110 <SUR>
 A;Cross-references: GB:J00265; NID:g186429; PIDN:AAA59172.1; PID:g386828
 R;Nicol, D.S.H.W.; Smith, L.F.
 Nature 187, 483-485, 1960
 A;Title: Amino-acid sequence of human insulin.
 A;Reference number: A93144
 A;Accession: A93144
 A;Molecule type: protein
 A;Residues: 25-54;90-110 <NIC>
 R;Oyer, P.E.; Cho, S.; Peterson, J.D.; Steiner, D.F.
 J. Biol. Chem. 246, 1375-1386, 1971
 A;Title: Studies on human proinsulin. Isolation and amino acid sequence of the human pancreatic C-peptide.
 A;Reference number: A92075; MUID:71116410; PMID:5101771
 A;Accession: A92075
 A;Molecule type: protein
 A;Residues: 57-87 <OYE>
 R;Ko, A.; Smyth, D.G.; Markussen, J.; Sundby, F.
 Eur. J. Biochem. 20, 190-199, 1971
 A;Title: Amino acid sequence of the C-peptide of human proinsulin.
 A;Reference number: A91186; MUID:71257722; PMID:5560404
 A;Accession: A91186
 A;Molecule type: protein
 A;Residues: 57-87 <KOA>
 R;Lucassen, A.M.; Julier, C.; Beressi, J.P.; Boitard, C.; Froguel, P.; Lathrop, M.; Bell, J.I.
 Nature Genet. 4, 305-310, 1993
 A;Title: Susceptibility to insulin dependent diabetes mellitus maps to a 4.1 kb segment of DNA spanning the insulin gene and associated VNTR.
 A;Reference number: I58114; MUID:93364428; PMID:8358440
 A;Accession: I58114
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-59,63-110 <RES>
 A;Cross-references: GB:L15440; NID:g307071; PIDN:AAA59179.1; PID:g307072
 R;Sieber, P.; Kamber, B.; Hartmann, A.; Joehl, A.; Riniker, B.; Rittel, W.
 Helv. Chim. Acta 57, 2617-2621, 1974

A;Title: Totalsynthese von Humaninsulin unter gezielter Bildung der Disulfidbindungen.
A;Reference number: A91636; MUID:75077277; PMID:4443293
A;Contents: annotation; synthesis
A;Note: disulfide-bonded human insulin was synthesized; the synthetic hormone was identical with the natural hormone in chemical and biological activities
A;Note: article in German with English abstract
R;Naithani, V.K.
Hoppe-Seyler's Z. Physiol. Chem. 354, 659-672, 1973
A;Title: The synthesis of C-peptide of human proinsulin.
A;Reference number: A91658; MUID:75040007; PMID:4803504
A;Contents: annotation; synthesis of residues 57-87
R;Geiger, R.; Jaeger, G.; Koenig, W.
Chem. Ber. 106, 2347-2352, 1973
A;Title: Synthesis of the complete sequence of human proinsulin C-peptide and its [Glu-9,Gln-11] analogue.
A;Reference number: A90914
A;Contents: annotation; synthesis of residues 57-87
R;Kaufmann, J.E.; Irminger, J.C.; Halban, P.A.
Biochem. J. 310, 869-874, 1995
A;Title: Sequence requirements for proinsulin processing at the B-chain/C-peptide junction.
A;Reference number: S58661; MUID:96013185; PMID:7575420
A;Contents: annotation; site-directed mutagenesis study of proteolytic processing
C;Genetics:
A;Gene: GDB:INS
A;Cross-references: GDB:119349; OMIM:176730
A;Map position: 11p15.5-11p15.5
A;Introns: 63/1
C;Superfamily: insulin
C;Keywords: hormone; pancreas
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-54/Domain: insulin chain B #status experimental <BCH>
F;25-54,90-110/Product: insulin #status experimental <MAT>
F;57-87/Domain: connecting C peptide #status experimental <CPEP>
F;90-110/Domain: insulin chain A #status experimental <ACH>
F;31-96,43-109,95-100/Disulfide bonds: #status experimental

Query Match 100.0%; Score 463; DB 1; Length 110;
Best Local Similarity 100.0%; Pred. No. 6.8e-43;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
          |||
Db      25 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 84

Qy      61 SLQKRGIVEQCCTSICSLYQLENYCN 86
          |||
Db      85 SLQKRGIVEQCCTSICSLYQLENYCN 110

```

RESULT 2

A42179

insulin precursor - chimpanzee

C;Species: Pan troglodytes (chimpanzee)

C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004

C;Accession: A42179; S22058
 R;Seino, S.; Bell, G.I.; Li, W.H.
 Mol. Biol. Evol. 9, 193-203, 1992
 A;Title: Sequences of primate insulin genes support the hypothesis of a slower rate of molecular evolution in humans and apes than in monkeys.
 A;Reference number: A42179; MUID:92219953; PMID:1560757
 A;Accession: A42179
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-110 <SEI>
 A;Cross-references: UNIPROT:P30410; EMBL:X61089; NID:g38251; PIDN:CAA43403.1; PID:g38252
 A;Note: sequence extracted from NCBI backbone (NCBIP:95067)
 C;Genetics:
 A;Introns: 63/1
 C;Superfamily: insulin

Query Match 100.0%; Score 463; DB 2; Length 110;
 Best Local Similarity 100.0%; Pred. No. 6.8e-43;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
          |||
Db      25 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 84

Qy      61 SLQKRGIVEQCCTSICSLYQLENYCN 86
          |||
Db      85 SLQKRGIVEQCCTSICSLYQLENYCN 110

```

RESULT 3

B42179

insulin precursor - green monkey

C;Species: Cercopithecus aethiops (green monkey, grivet)

C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004

C;Accession: B42179; A05232; S16494; S22056

R;Seino, S.; Bell, G.I.; Li, W.H.

Mol. Biol. Evol. 9, 193-203, 1992

A;Title: Sequences of primate insulin genes support the hypothesis of a slower rate of molecular evolution in humans and apes than in monkeys.

A;Reference number: A42179; MUID:92219953; PMID:1560757

A;Accession: B42179

A;Molecule type: DNA

A;Residues: 1-110 <SEI>

A;Cross-references: UNIPROT:P30407; EMBL:X61092; NID:g22808; PIDN:CAA43405.1; PID:g22809

A;Note: sequence extracted from NCBI backbone (NCBIN:95185, NCBIP:95194)

R;Peterson, J.D.; Nehrlich, S.; Oyer, P.E.; Steiner, D.F.

J. Biol. Chem. 247, 4866-4871, 1972

A;Title: Determination of the amino acid sequence of the monkey, sheep, and dog proinsulin C-peptides by a semi-micro Edman degradation procedure.

A;Reference number: A92111; MUID:72258016; PMID:4626369

A;Accession: A05232

A;Molecule type: protein

A;Residues: 57-87 <PET>

C;Genetics:

A;Introns: 63/1

C;Superfamily: insulin
C;Keywords: hormone; pancreas
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-54/Domain: insulin chain B #status predicted <BCH>
F;25-54,90-110/Product: insulin #status predicted <MAT>
F;57-87/Domain: connecting peptide #status experimental <CPEP>
F;90-110/Domain: insulin chain A #status predicted <ACH>
F;31-96,43-109,95-100/Disulfide bonds: #status predicted

Query Match 98.5%; Score 456; DB 2; Length 110;
Best Local Similarity 98.8%; Pred. No. 3.9e-42;
Matches 85; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
          |||
Db      25 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDPQVGQVELGGGPGAGSLQPLALEG 84

Qy      61 SLQKRGIVEQCCTSICSLYQLENYCN 86
          |||
Db      85 SLQKRGIVEQCCTSICSLYQLENYCN 110
```

RESULT 4

JQ0178

insulin precursor - crab-eating macaque

C;Species: *Macaca fascicularis* (crab-eating macaque)

C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004

C;Accession: JQ0178

R;Wetekam, W.; Groneberg, J.; Leineweber, M.; Wengenmayer, F.; Winnacker, E.L.
Gene 19, 179-183, 1982

A;Title: The nucleotide sequence of cDNA coding for preproinsulin from the
primate *Macaca fascicularis*.

A;Reference number: JQ0178; MUID:83080474; PMID:6184262

A;Accession: JQ0178

A;Molecule type: mRNA

A;Residues: 1-110 <WET>

A;Cross-references: UNIPROT:P30406; GB:J00336; NID:g342121; PIDN:AAA36849.1;
PID:g342122

C;Superfamily: insulin

F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-54,90-110/Product: insulin #status predicted <MAT>
F;25-54/Domain: insulin chain B #status predicted <BCH>
F;55-89/Domain: insulin connecting C peptide #status predicted <CPT>
F;90-110/Domain: insulin chain A #status predicted <ACH>
F;31-96,43-109,95-100/Disulfide bonds: #status predicted

Query Match 98.5%; Score 456; DB 2; Length 110;
Best Local Similarity 98.8%; Pred. No. 3.9e-42;
Matches 85; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
          |||
Db      25 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDPQVGQVELGGGPGAGSLQPLALEG 84

Qy      61 SLQKRGIVEQCCTSICSLYQLENYCN 86
          |||
Db      85 SLQKRGIVEQCCTSICSLYQLENYCN 110
```

R; Kwok, S.C.M.; Chan, S.J.; Steiner, D.F.

A;Reference number: A90082

A:Reference number: A48172; MUID:90372989; PMID:2397023

A;Accession: B48172
 A;Status: not compared with conceptual translation
 A;Molecule type: DNA
 A;Residues: 1-110 <SAW>
 R;Linde, S.; Nielsen, J.H.; Hansen, B.; Welinder, B.S.
 J. Chromatogr. 462, 243-254, 1989
 A;Title: Reversed-phase high-performance liquid chromatographic analyses of insulin biosynthesis in isolated rat and mouse islets.
 A;Reference number: A61012; MUID:89292078; PMID:2661585
 A;Accession: A61012
 A;Molecule type: protein
 A;Residues: 57-87 <LIN>
 R;Buenzli, H.F.; Glatthaar, B.; Kunz, P.; Muelhaupt, E.; Humbel, R.E.
 Hoppe-Seyler's Z. Physiol. Chem. 353, 451-458, 1972
 A;Title: Amino acid sequence of the two insulins from mouse (*Mus musculus*).
 A;Reference number: A01592; MUID:72189455; PMID:5063718
 A;Accession: B01592
 A;Molecule type: protein
 A;Residues: 25-54;90-110 <BUE>
 C;Genetics:
 A;Introns: 63/1
 C;Superfamily: insulin
 C;Keywords: hormone; pancreas
 F;1-24/Domain: signal sequence #status predicted <SIG>
 F;25-54/Domain: insulin chain B #status experimental <BCH>
 F;25-54,90-110/Product: insulin #status experimental <MAT>
 F;57-87/Domain: connecting peptide #status experimental <CPEP>
 F;90-110/Domain: insulin chain A #status experimental <ACH>
 F;31-96,43-109,95-100/Disulfide bonds: #status predicted

Query Match 85.1%; Score 394; DB 1; Length 110;
 Best Local Similarity 84.9%; Pred. No. 1.9e-35;
 Matches 73; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

```

Qy      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      25 FVKQHLCGSHLVEALYLVCGERGFFYTPMSRREVEDPQVAQLLELGGGPGAGDLQTLALEV 84

Qy      61 SLQKRGIVEQCCTSICSLYQLENYCN 86
      : |||||:||||| ||||| |||||
Db      85 AQQKRGIVDQCCTSICSLYQLENYCN 110
  
```

RESULT 9

IPRT2
 insulin 2 precursor - rat
 C;Species: *Rattus norvegicus* (Norway rat)
 C;Date: 23-Oct-1981 #sequence_revision 23-Oct-1981 #text_change 09-Jul-2004
 C;Accession: B90789; B94231; C92120; I64880; A01590; B92120
 R;Lomedico, P.; Rosenthal, N.; Efstratiadis, A.; Gilbert, W.; Kolodner, R.; Tizard, R.
 Cell 18, 545-558, 1979
 A;Title: The structure and evolution of the two nonallelic rat preproinsulin genes.
 A;Reference number: A90789; MUID:80045035; PMID:498284
 A;Accession: B90789
 A;Molecule type: DNA

A;Residues: 1-110 <LOM>
A;Cross-references: UNIPROT:P01323; GB:J00748; NID:g204958; PIDN:AAA41443.1; PID:g204959
R;Steiner, D.F.; Clark, J.L.; Nolan, C.; Rubenstein, A.H.; Margoliash, E.; Aten, B.; Oyer, P.E.
Recent Prog. Horm. Res. 25, 207-282, 1969
A;Title: Proinsulin and the biosynthesis of insulin.
A;Reference number: A94231; MUID:70067613; PMID:4311938
A;Accession: B94231
A;Molecule type: protein
A;Residues: 25-54;90-110 <STE>
R;Tager, H.S.; Steiner, D.F.
J. Biol. Chem. 247, 7936-7940, 1972
A;Title: Primary structures of the proinsulin connecting peptides of the rat and horse.
A;Reference number: A92120; MUID:73061498; PMID:4640931
A;Accession: C92120
A;Molecule type: protein
A;Residues: 57-87 <TAG>
R;Lomedico, P.T.; Rosenthal, N.; Kolodner, R.; Efstratiadis, A.; Gilbert, W.
Ann. N. Y. Acad. Sci. 343, 425-432, 1980
A;Title: The structure of rat preproinsulin genes.
A;Reference number: I51945; MUID:80240379; PMID:6249167
A;Accession: I64880
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-110 <RES>
A;Cross-references: GB:M25585; NID:g204950; PIDN:AAA41440.1; PID:g204952
C;Genetics:
A;Gene: INS2
A;Introns: 63/1
C;Superfamily: insulin
C;Keywords: hormone; pancreas
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-54/Domain: insulin chain B #status experimental <BCH>
F;25-54,90-110/Product: insulin #status experimental <MAT>
F;57-87/Domain: connecting peptide #status experimental <CPEP>
F;90-110/Domain: insulin chain A #status experimental <ACH>
F;31-96,43-109,95-100/Disulfide bonds: #status experimental

Query Match 85.1%; Score 394; DB 1; Length 110;
Best Local Similarity 84.9%; Pred. No. 1.9e-35;
Matches 73; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

```
Qy      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
        || |||||
Db      25 FVKQHLCGSHLVEALYLVCGERGFFYTPMSRREVDPQVAQLELGGGPGAGDLQTLALEV 84
        || |||||

Qy      61 SLQKRGIVEQCCTSICSLYQLENYCN 86
        : |||||
Db      85 ARQKRGIVDQCCTSICSLYQLENYCN 110
```

RESULT 10

A39883

insulin precursor - douroucouli

C;Species: Aotus trivirgatus (douroucouli, night monkey, owl monkey)

C;Date: 27-Nov-1991 #sequence_revision 27-Nov-1991 #text_change 09-Jul-2004
C;Accession: A39883
R;Seino, S.; Steiner, D.F.; Bell, G.I.
Proc. Natl. Acad. Sci. U.S.A. 84, 7423-7427, 1987
A;Title: Sequence of a New World primate insulin having low biological potency
and immunoreactivity.
A;Reference number: A39883; MUID:88041119; PMID:3118367
A;Accession: A39883
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-108 <SEI>
A;Cross-references: UNIPROT:P10604; GB:J02989; NID:gl76555; PIDN:AAA35374.1;
PID:gl76556
C;Superfamily: insulin

RESULT 11

I48166
insulin precursor - golden hamster
C;Species: *Mesocricetus auratus* (golden hamster)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 16-Jul-1999
C;Accession: I48166
R;Bell, G.I.; Sanchez-Pescador, R.
Diabetes 33, 297-300, 1984
A;Title: Sequence of a cDNA encoding Syrian hamster preproinsulin.
A;Reference number: I48166; MUID:84133036; PMID:6365663
A;Accession: I48166
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-110 <RES>
A;Cross-references: GB:M26328; NID:g191420; PIDN:AAA37089.1; PID:g305360
C;Superfamily: insulin

RESULT 12

IPRT1

insulin 1 precursor - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 23-Oct-1981 #sequence_revision 23-Oct-1981 #text_change 09-Jul-2004

C;Accession: A90788; A90789; A94231; B92120; I51945; A01589

R;Cordell, B.; Bell, G.; Tischer, E.; DeNoto, F.M.; Ullrich, A.; Pictet, R.; Rutter, W.J.; Goodman, H.M.

Cell 18, 533-543, 1979

A;Title: Isolation and characterization of a cloned rat insulin gene.

A;Reference number: A90788; MUID:80045034; PMID:498283

A;Accession: A90788

A;Molecule type: DNA

A;Residues: 1-110 <COR>

A;Cross-references: UNIPROT:P01322; GB:J00747; NID:g204956; PIDN:AAA41442.1; PID:g204957

R;Lomedico, P.; Rosenthal, N.; Efstratiadis, A.; Gilbert, W.; Kolodner, R.; Tizard, R.

Cell 18, 545-558, 1979

A;Title: The structure and evolution of the two nonallelic rat preproinsulin genes.

A;Reference number: A90789; MUID:80045035; PMID:498284

A;Accession: A90789

A;Molecule type: DNA

A;Residues: 1-110 <LOM>

A;Cross-references: GB:J00747; NID:g204956; PIDN:AAA41442.1; PID:g204957

R;Steiner, D.F.; Clark, J.L.; Nolan, C.; Rubenstein, A.H.; Margoliash, E.; Aten, B.; Oyer, P.E.

Recent Prog. Horm. Res. 25, 207-282, 1969

A;Title: Proinsulin and the biosynthesis of insulin.

A;Reference number: A94231; MUID:70067613; PMID:4311938

A;Accession: A94231

A;Molecule type: protein

A;Residues: 25-54;90-110 <STE>

R;Tager, H.S.; Steiner, D.F.

J. Biol. Chem. 247, 7936-7940, 1972

A;Title: Primary structures of the proinsulin connecting peptides of the rat and horse.

A;Reference number: A92120; MUID:73061498; PMID:4640931

A;Accession: B92120

A;Molecule type: protein

A;Residues: 57-87 <TAG>

R;Lomedico, P.T.; Rosenthal, N.; Kolodner, R.; Efstratiadis, A.; Gilbert, W. Ann. N. Y. Acad. Sci. 343, 425-432, 1980

A;Title: The structure of rat preproinsulin genes.

A;Reference number: I51945; MUID:80240379; PMID:6249167

A;Accession: I51945

A;Status: translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-110 <RES>

A;Cross-references: GB:M25584; NID:g204947; PIDN:AAA41439.1; PID:g204948

C;Genetics:

A;Gene: INS1

C;Superfamily: insulin

C;Keywords: hormone; pancreas

F;1-24/Domain: signal sequence #status predicted <SIG>

F;25-54/Domain: insulin chain B #status experimental <BCH>
F;25-54,90-110/Product: insulin #status experimental <MAT>
F;57-87/Domain: connecting peptide #status experimental <CPEP>
F;90-110/Domain: insulin chain A #status experimental <ACH>
F;31-96,43-109,95-100/Disulfide bonds: #status experimental

Query Match 83.2%; Score 385; DB 1; Length 110;
Best Local Similarity 83.7%; Pred. No. 1.8e-34;
Matches 72; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
|| ||||| |||||||||||||||||:|||| || || |:||||| || || |||||
Db 25 FVKQHLCGPHLVEALYLVCGERGFFYTPKSRREVEDPQVPQLELGGGPEAGDLQTLALEV 84

Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 86
: |||||||:|||||||||||||||||
Db 85 ARQKRGIVDQCCTSICSLYQLENYCN 110

RESULT 13

IPPG

insulin precursor - pig

C;Species: Sus scrofa domestica (domestic pig)

C;Date: 22-Jun-1981 #sequence_revision 22-Jun-1981 #text_change 16-Jul-1999

C;Accession: A01583; A94572; S16492; A60835; B60835

R;Chance, R.E.; Ellis, R.M.; Bromer, W.W.

Science 161, 165-167, 1968

A;Title: Porcine proinsulin: characterization and amino acid sequence.

A;Reference number: A94240; MUID:68286485; PMID:5657063

A;Accession: A01583

A;Molecule type: protein

A;Residues: 1-34,'Q',36-84 <CHA>

R;Chance, R.E.

submitted to the Atlas, July 1970

A;Reference number: A94572

A;Accession: A94572

A;Molecule type: protein

A;Residues: 1-84 <CH2>

R;Brown, H.; Sanger, F.; Kitai, R.

Biochem. J. 60, 556-565, 1955

A;Title: The structure of pig and sheep insulins.

A;Reference number: A90344

A;Accession: S16492

A;Molecule type: protein

A;Residues: 1-30;31-51 <BRO>

R;Snel, L.; Damgaard, U.

Horm. Metab. Res. 20, 476-480, 1988

A;Title: Proinsulin heterogeneity in pigs.

A;Reference number: A60835; MUID:89032178; PMID:3181865

A;Accession: A60835

A;Molecule type: protein

A;Residues: 33-38,40-62 <SNE>

A;Note: the authors report the characterization of a connecting peptide variant lacking Ala-39

A;Accession: B60835

A;Molecule type: protein

A;Residues: 33-62 <SN2>

R;Blundell, T.; Dodson, G.; Hodgkin, D.; Mercola, D.
 Adv. Protein Chem. 26, 279-402, 1972
 A;Title: Insulin. the structure in the crystal and its reflection in chemistry and biology.
 A;Reference number: A90017
 A;Contents: annotation; X-ray crystallography, 1.9 angstroms
 C;Superfamily: insulin
 C;Keywords: hormone; pancreas
 F;1-30/Domain: insulin chain B #status experimental <BCH>
 F;1-30,64-84/Product: insulin #status experimental <MAT>
 F;33-63/Domain: connecting peptide #status experimental <CPEP>
 F;64-84/Domain: insulin chain A #status experimental <ACH>
 F;7-70,19-83,69-74/Disulfide bonds: #status experimental

Query Match 82.7%; Score 383; DB 1; Length 84;
 Best Local Similarity 86.0%; Pred. No. 2.3e-34;
 Matches 74; Conservative 1; Mismatches 9; Indels 2; Gaps 1;

```
Qy      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
          |||||: || ||| |||
Db      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKARREAENPQAGAVELGG--GLGGLQALALEG 58

Qy      61 SLQKRGIVEQCCTSICSLYQLENYCN 86
          |||||
Db      59 PPQKRGIVEQCCTSICSLYQLENYCN 84
```

RESULT 14

IPBO

insulin precursor - bovine

C;Species: Bos primigenius taurus (cattle)

C;Date: 24-Apr-1984 #sequence_revision 22-Apr-1995 #text_change 09-Jul-2004

C;Accession: A40909; A92080; A92074; A91185; A90342; A90341; S48184; S48185; S46258; A01585

R;D'Agostino, J.; Younes, M.A.; White, J.W.; Besch, P.K.; Field, J.B.; Frazier, M.L.

Mol. Endocrinol. 1, 327-331, 1987

A;Title: Cloning and nucleotide sequence analysis of complementary deoxyribonucleic acid for bovine preproinsulin.

A;Reference number: A40909; MUID:88288209; PMID:2456452

A;Accession: A40909

A;Molecule type: mRNA

A;Residues: 1-105 <DAA>

A;Cross-references: UNIPROT:P01317; GB:M54979; NID:g163578; PIDN:AAA30722.1; PID:g163579

A;Experimental source: fetal pancreas

R;Nolan, C.; Margoliash, E.; Peterson, J.D.; Steiner, D.F.

J. Biol. Chem. 246, 2780-2795, 1971

A;Title: The structure of bovine proinsulin.

A;Reference number: A92080; MUID:71166442; PMID:4928892

A;Accession: A92080

A;Molecule type: protein

A;Residues: 25-105 <NOL>

R;Steiner, D.F.; Cho, S.; Oyer, P.E.; Terris, S.; Peterson, J.D.; Rubenstein, A.H.

J. Biol. Chem. 246, 1365-1374, 1971

A;Title: Isolation and characterization of proinsulin C-peptide from bovine pancreas.
 A;Reference number: A92074; MUID:71116409; PMID:5545080
 A;Accession: A92074
 A;Molecule type: protein
 A;Residues: 57-82 <STE>
 R;Salokangas, A.; Smyth, D.G.; Markussen, J.; Sundby, F.
 Eur. J. Biochem. 20, 183-189, 1971
 A;Title: Bovine proinsulin: amino acid sequence of the C-peptide isolated from pancreas.
 A;Reference number: A91185; MUID:71257721; PMID:5105368
 A;Accession: A91185
 A;Molecule type: protein
 A;Residues: 57-82 <SAL>
 R;Sanger, F.; Thompson, E.O.P.
 Biochem. J. 53, 366-374, 1953
 A;Title: The amino-acid sequence in the glycyl chain of insulin. 2. The investigation of peptides from enzymic hydrolysates.
 A;Reference number: A90342
 A;Accession: A90342
 A;Molecule type: protein
 A;Residues: 85-105 <SAN>
 R;Sanger, F.; Tuppy, H.
 Biochem. J. 49, 481-490, 1951
 A;Title: The amino-acid sequence in the phenylalanyl chain of insulin. 2. The investigation of peptides from enzymic hydrolysates.
 A;Reference number: A90341
 A;Accession: A90341
 A;Molecule type: protein
 A;Residues: 25-54 <SA2>
 R;Cheng, R.; Kawakishi, S.
 Eur. J. Biochem. 223, 759-764, 1994
 A;Title: Site-specific oxidation of histidine residues in glycated insulin mediated by Cu(2+).
 A;Reference number: S48184; MUID:94333378; PMID:8055951
 A;Accession: S48184
 A;Molecule type: protein
 A;Residues: 85-105 <CHE>
 A;Accession: S48185
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 25-30,'X',32-42,'X',44-54 <CH2>
 R;Ryle, A.P.; Sanger, F.; Smith, L.F.; Kitai, R.
 Biochem. J. 60, 541-556, 1955
 A;Title: The disulphide bonds of insulin.
 A;Reference number: A90343
 A;Contents: annotation; amides; disulfides
 R;Wenzel, T.; Eckerskorn, C.; Lottspeich, F.; Baumeister, W.
 FEBS Lett. 349, 205-209, 1994
 A;Title: Existence of a molecular ruler in proteasomes suggested by analysis of degradation products.
 A;Reference number: S46258; MUID:94326921; PMID:8050567
 A;Accession: S46258
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 25-54 <WEN>
 C;Superfamily: insulin

C;Keywords: hormone; pancreas
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-54/Domain: insulin chain B #status experimental <BCH>
F;25-54,85-105/Product: insulin #status experimental <MAT>
F;57-82/Domain: connecting peptide #status experimental <CPEP>
F;85-105/Domain: insulin chain A #status experimental <ACH>
F;31-91,43-104,90-95/Disulfide bonds: #status experimental

Query Match 79.2%; Score 366.5; DB 1; Length 105;
Best Local Similarity 80.2%; Pred. No. 1.7e-32;
Matches 69; Conservative 2; Mismatches 10; Indels 5; Gaps 1;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
|||||
Db 25 FVNQHLCGSHLVEALYLVCGERGFFYTPKARREVEGPQVGALELAGGPGAG-----GLEG 79
Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 86
|||||
Db 80 PPQKRGIVEQCCASVCSLYQLENYCN 105

RESULT 15

INMS1

insulin 1 precursor - mouse

C;Species: Mus musculus (house mouse)

C;Date: 24-Apr-1984 #sequence_revision 14-Jul-1994 #text_change 09-Jul-2004

C;Accession: B26342; A48172; A01592; B61012

R;Wentworth, B.M.; Schaefer, I.M.; Villa-Komaroff, L.; Chirgwin, J.M.

J. Mol. Evol. 23, 305-312, 1986

A;Title: Characterization of the two nonallelic genes encoding mouse preproinsulin.

A;Reference number: A92965; MUID:87169768; PMID:3104603

A;Accession: B26342

A;Molecule type: DNA

A;Residues: 1-108 <WEN>

A;Cross-references: UNIPROT:P01325; GB:X04725; NID:g52712; PIDN:CAA28434.1;
PID:g52713

R;Sawa, T.; Ohgaku, S.; Morioka, H.; Yano, S.

J. Mol. Endocrinol. 5, 61-67, 1990

A;Title: Molecular cloning and DNA sequence analysis of preproinsulin genes in the NON mouse, an animal model of human non-obese, non-insulin-dependent diabetes mellitus.

A;Reference number: A48172; MUID:90372989; PMID:2397023

A;Accession: A48172

A;Status: not compared with conceptual translation

A;Molecule type: DNA

A;Residues: 1-108 <SAW>

R;Buenzli, H.F.; Glatthaar, B.; Kunz, P.; Muelhaupt, E.; Humbel, R.E.

Hoppe-Seyler's Z. Physiol. Chem. 353, 451-458, 1972

A;Title: Amino acid sequence of the two insulins from mouse (Mus musculus).

A;Reference number: A01592; MUID:72189455; PMID:5063718

A;Accession: A01592

A;Molecule type: protein

A;Residues: 25-54;88-108 <BUE>

R;Linde, S.; Nielsen, J.H.; Hansen, B.; Welinder, B.S.

J. Chromatogr. 462, 243-254, 1989

A;Title: Reversed-phase high-performance liquid chromatographic analyses of insulin biosynthesis in isolated rat and mouse islets.

A;Reference number: A61012; MUID:89292078; PMID:2661585

A;Accession: B61012

A;Molecule type: protein

A;Residues: 57-85 <LIN>

C;Superfamily: insulin

C;Keywords: hormone; pancreas

F;1-24/Domain: signal sequence #status predicted <SIG>

F;25-54/Domain: insulin chain B #status experimental <BCH>

F;25-54,88-108/Product: insulin #status experimental <MAT>

F;57-85/Domain: connecting peptide #status experimental <CPEP>

F;88-108/Domain: insulin chain A #status experimental <ACH>

F;31-94,43-107,93-98/Disulfide bonds: #status predicted

Query Match 79.0%; Score 366; DB 1; Length 108;

Best Local Similarity 81.4%; Pred. No. 2e-32;

Matches 70; Conservative 4; Mismatches 10; Indels 2; Gaps 1;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG .60
|| ||||| |||||||||||||||||||||:|||| || || |:|||| | | || |||||

Db 25 FVKQHLCGPHLVEALYLVCGERGFFYTPKSRREVEDPQVEQLELGGSP--GDLQTLALEV 82

Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 86
: |||||||:|||||||||||||||||

Db 83 ARQKRGIVDQCCTSICSLYQLENYCN 108

Search completed: February 11, 2005, 18:24:35

Job time : 17.3432 secs

OM protein - protein search, using sw model

Run on: February 11, 2005, 18:23:02 ; Search time 62.5166 Seconds
(without alignments)
449.487 Million cell updates/sec

Title: US-10-054-873-4
Perfect score: 463
Sequence: 1 FVNQHLCGSHLVEALYLVCG.....IVEQCCTSICSLYQLENYCN 86

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1376875 seqs, 326749119 residues

Total number of hits satisfying chosen parameters: 1376875

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|-------------------|-------------------|
| 1 | 463 | 100.0 | 86 | 9 | US-09-878-380-1 | Sequence 1, Appli |
| 2 | 463 | 100.0 | 86 | 10 | US-09-858-935B-4 | Sequence 4, Appli |
| 3 | 463 | 100.0 | 86 | 13 | US-10-028-410-2 | Sequence 2, Appli |
| 4 | 463 | 100.0 | 86 | 13 | US-10-054-873-4 | Sequence 4, Appli |
| 5 | 463 | 100.0 | 86 | 14 | US-10-444-326-2 | Sequence 2, Appli |
| 6 | 463 | 100.0 | 86 | 15 | US-10-271-869-4 | Sequence 4, Appli |
| 7 | 463 | 100.0 | 86 | 15 | US-10-444-262-2 | Sequence 2, Appli |
| 8 | 463 | 100.0 | 86 | 15 | US-10-444-649-2 | Sequence 2, Appli |
| 9 | 463 | 100.0 | 86 | 15 | US-10-444-701-2 | Sequence 2, Appli |
| 10 | 463 | 100.0 | 86 | 17 | US-10-760-928-2 | Sequence 2, Appli |
| 11 | 463 | 100.0 | 96 | 9 | US-09-947-563-4 | Sequence 4, Appli |
| 12 | 463 | 100.0 | 110 | 9 | US-09-205-658-125 | Sequence 125, App |
| 13 | 463 | 100.0 | 110 | 9 | US-09-815-229-3 | Sequence 3, Appli |
| 14 | 463 | 100.0 | 110 | 9 | US-09-804-409A-9 | Sequence 9, Appli |
| 15 | 463 | 100.0 | 110 | 10 | US-09-969-748C-6 | Sequence 6, Appli |
| 16 | 463 | 100.0 | 110 | 10 | US-09-963-693-125 | Sequence 125, App |
| 17 | 463 | 100.0 | 110 | 14 | US-10-038-686-1 | Sequence 1, Appli |
| 18 | 463 | 100.0 | 110 | 14 | US-10-328-813-2 | Sequence 2, Appli |
| 19 | 463 | 100.0 | 110 | 15 | US-10-383-285-2 | Sequence 2, Appli |
| 20 | 463 | 100.0 | 110 | 15 | US-10-346-563-2 | Sequence 2, Appli |
| 21 | 463 | 100.0 | 110 | 15 | US-10-321-717-2 | Sequence 2, Appli |
| 22 | 463 | 100.0 | 110 | 15 | US-10-411-037-44 | Sequence 44, Appl |
| 23 | 463 | 100.0 | 110 | 15 | US-10-411-026-44 | Sequence 44, Appl |
| 24 | 463 | 100.0 | 110 | 15 | US-10-410-962-44 | Sequence 44, Appl |
| 25 | 463 | 100.0 | 110 | 15 | US-10-411-049-44 | Sequence 44, Appl |
| 26 | 463 | 100.0 | 110 | 15 | US-10-700-725-20 | Sequence 20, Appl |
| 27 | 463 | 100.0 | 110 | 16 | US-10-410-930-44 | Sequence 44, Appl |
| 28 | 463 | 100.0 | 110 | 16 | US-10-410-997-44 | Sequence 44, Appl |
| 29 | 463 | 100.0 | 110 | 16 | US-10-411-012-44 | Sequence 44, Appl |
| 30 | 463 | 100.0 | 110 | 16 | US-10-287-994-44 | Sequence 44, Appl |
| 31 | 463 | 100.0 | 110 | 16 | US-10-740-098-3 | Sequence 3, Appli |
| 32 | 463 | 100.0 | 110 | 16 | US-10-410-913-44 | Sequence 44, Appl |
| 33 | 463 | 100.0 | 117 | 9 | US-09-280-030-63 | Sequence 63, Appl |
| 34 | 463 | 100.0 | 130 | 9 | US-09-280-030-62 | Sequence 62, Appl |
| 35 | 457 | 98.7 | 96 | 9 | US-09-947-563-5 | Sequence 5, Appli |
| 36 | 442 | 95.5 | 110 | 16 | US-10-419-539-5 | Sequence 5, Appli |
| 37 | 438.5 | 94.7 | 124 | 15 | US-10-221-677-24 | Sequence 24, Appl |
| 38 | 388 | 83.8 | 86 | 17 | US-10-760-928-1 | Sequence 1, Appli |
| 39 | 383 | 82.7 | 84 | 17 | US-10-760-928-3 | Sequence 3, Appli |
| 40 | 366.5 | 79.2 | 81 | 17 | US-10-760-928-4 | Sequence 4, Appli |
| 41 | 306 | 66.1 | 166 | 9 | US-09-925-297-805 | Sequence 805, App |
| 42 | 300 | 64.8 | 56 | 9 | US-09-815-229-10 | Sequence 10, Appl |
| 43 | 300 | 64.8 | 56 | 16 | US-10-740-098-10 | Sequence 10, Appl |
| 44 | 285 | 61.6 | 54 | 9 | US-09-815-229-13 | Sequence 13, Appl |
| 45 | 285 | 61.6 | 54 | 16 | US-10-740-098-13 | Sequence 13, Appl |

ALIGNMENTS

RESULT 1

US-09-878-380-1

; Sequence 1, Application US/09878380

; Patent No. US20020160435A1

```
; GENERAL INFORMATION:
; APPLICANT: Fujirebio Inc.
; APPLICANT: KITAJIMA, Sachiko
; APPLICANT: KURANO, Yoshihiro
; APPLICANT: NAKATSUBO, Kaoru
; APPLICANT: NISHIZONO, Isao
; TITLE OF INVENTION: Immunoassay For Measuring Human C-Peptide and Kit
Therefor
; FILE REFERENCE: 0760-0291P
; CURRENT APPLICATION NUMBER: US/09/878,380
; CURRENT FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: JP 2000-174691
; PRIOR FILING DATE: 2000-06-12
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-878-380-1
```

```
Query Match          100.0%; Score 463; DB 9; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.3e-44;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
          |||
Db      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60

Qy      61 SLQKRGIVEQCCTSICSLYQLENYCN 86
          |||
Db      61 SLQKRGIVEQCCTSICSLYQLENYCN 86
```

RESULT 2

```
US-09-858-935B-4
; Sequence 4, Application US/09858935B
; Publication No. US20030069177A1
; GENERAL INFORMATION:
; APPLICANT: Dubaquie, Yves
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Lowman, Henry B.
; TITLE OF INVENTION: METHOD FOR TREATING CARTILAGE DISORDERS
; FILE REFERENCE: P1794R1
; CURRENT APPLICATION NUMBER: US/09/858,935B
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: US 60/248,985
; PRIOR FILING DATE: 2000-11-15
; PRIOR APPLICATION NUMBER: US 60/204,490
; PRIOR FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 153
; SEQ ID NO 4
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-858-935B-4
```

Query Match 100.0%; Score 463; DB 10; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.3e-44;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
|
Db 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60

Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 86
|
Db 61 SLQKRGIVEQCCTSICSLYQLENYCN 86

RESULT 3

US-10-028-410-2
; Sequence 2, Application US/10028410
; Publication No. US20020160955A1
; GENERAL INFORMATION:
; APPLICANT: Dubaquier, Yves
; APPLICANT: Lowman, Henry
; TITLE OF INVENTION: PROTEIN VARIANTS
; FILE REFERENCE: P1712R1-1
; CURRENT APPLICATION NUMBER: US/10/028,410
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US/09/477,924
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 2
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-028-410-2

Query Match 100.0%; Score 463; DB 13; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.3e-44;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
|
Db 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60

Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 86
|
Db 61 SLQKRGIVEQCCTSICSLYQLENYCN 86

RESULT 4

US-10-054-873-4
; Sequence 4, Application US/10054873
; Publication No. US20020164712A1
; GENERAL INFORMATION:
; APPLICANT: Gan, Zhong Ru
; TITLE OF INVENTION: Chimeric Protein Containing an
; Intramolecular Chaperone-Like Sequence
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP

```

; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/054,873
; FILING DATE: 22-Jan-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/CN98/00052
; FILING DATE: 31-MAR-1998
; APPLICATION NUMBER: US 09/423,100
; FILING DATE: 11-DEC-2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Mycroft, Frank J
; REGISTRATION NUMBER: 46,946
; REFERENCE/DOCKET NUMBER: 020167-000130US
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 86 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-054-873-4

```

```

Query Match          100.0%; Score 463; DB 13; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.3e-44;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60

Qy      61 SLQKRGIVEQCCTSICSLYQLENYCN 86
        ||||||||||||||||||
Db      61 SLQKRGIVEQCCTSICSLYQLENYCN 86

```

```

RESULT 5
US-10-444-326-2
; Sequence 2, Application US/10444326
; Publication No. US20030191065A1
; GENERAL INFORMATION:
; APPLICANT: Dubaquié, Yves
; APPLICANT: Lowman, Henry
; TITLE OF INVENTION: PROTEIN VARIANTS
; FILE REFERENCE: P1712R1
; CURRENT APPLICATION NUMBER: US/10/444,326
; CURRENT FILING DATE: 2003-05-22

```

```

Query Match          100.0%;  Score 463;  DB 14;  Length 86;
Best Local Similarity 100.0%;  Pred. No. 1.3e-44;
Matches      86;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

Qy          1  FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
             |||
Db          1  FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60

Qy          61  SLQKRGIVEQCCTSICSLYQLENYCN 86
             |||
Db          61  SLQKRGIVEQCCTSICSLYQLENYCN 86

```

US-10-271-869-4

Query Match 100.0%; Score 463; DB 15; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.3e-44;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
|||||

Db 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60

Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 86
|||||
Db 61 SLQKRGIVEQCCTSICSLYQLENYCN 86

RESULT 7

US-10-444-262-2
; Sequence 2, Application US/10444262
; Publication No. US20040023883A1
; GENERAL INFORMATION:
; APPLICANT: Dubaquie, Yves
; APPLICANT: Lowman, Henry
; TITLE OF INVENTION: PROTEIN VARIANTS
; FILE REFERENCE: P1712R1
; CURRENT APPLICATION NUMBER: US/10/444,262
; CURRENT FILING DATE: 2003-05-22
; PRIOR APPLICATION NUMBER: US/09/724,478
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US/09/477,923
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 2
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-444-262-2

Query Match 100.0%; Score 463; DB 15; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.3e-44;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
|||||
Db 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60

Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 86
|||||
Db 61 SLQKRGIVEQCCTSICSLYQLENYCN 86

RESULT 8

US-10-444-649-2
; Sequence 2, Application US/10444649
; Publication No. US20040033951A1
; GENERAL INFORMATION:
; APPLICANT: Dubaquie, Yves
; APPLICANT: Lowman, Henry
; TITLE OF INVENTION: PROTEIN VARIANTS
; FILE REFERENCE: P1712R1
; CURRENT APPLICATION NUMBER: US/10/444,649
; CURRENT FILING DATE: 2003-05-22
; PRIOR APPLICATION NUMBER: US/09/724,479
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US/09/477,923
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 2

; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-444-649-2

Query Match 100.0%; Score 463; DB 15; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.3e-44;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
|
Db 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60

Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 86
|
Db 61 SLQKRGIVEQCCTSICSLYQLENYCN 86

RESULT 9

US-10-444-701-2
; Sequence 2, Application US/10444701
; Publication No. US20040033952A1
; GENERAL INFORMATION:
; APPLICANT: Dubaquier, Yves
; APPLICANT: Lowman, Henry
; TITLE OF INVENTION: PROTEIN VARIANTS
; FILE REFERENCE: P1712R1
; CURRENT APPLICATION NUMBER: US/10/444,701
; CURRENT FILING DATE: 2003-05-22
; PRIOR APPLICATION NUMBER: US/09/723,866
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US/09/477,923
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 2
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-444-701-2

Query Match 100.0%; Score 463; DB 15; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.3e-44;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
|
Db 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60

Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 86
|
Db 61 SLQKRGIVEQCCTSICSLYQLENYCN 86

RESULT 10

US-10-760-928-2
; Sequence 2, Application US/10760928
; Publication No. US20050026826A1

```
; GENERAL INFORMATION:
; APPLICANT: HOENIG, MARGARETHE
; TITLE OF INVENTION: FELINE PROINSULIN, INSULIN AND CONSTITUENT PEPTIDES
; FILE REFERENCE: 235.00520101
; CURRENT APPLICATION NUMBER: US/10/760,928
; CURRENT FILING DATE: 2004-01-20
; PRIOR APPLICATION NUMBER: 60/440,964
; PRIOR FILING DATE: 2003-01-17
; PRIOR APPLICATION NUMBER: 60/444,009
; PRIOR FILING DATE: 2003-01-31
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 2
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-760-928-2
```

```
Query Match          100.0%; Score 463; DB 17; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.3e-44;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60

QY      61 SLQKRGIVEQCCTSICSLYQLENYCN 86
          ||||||||||||||||||
Db      61 SLQKRGIVEQCCTSICSLYQLENYCN 86
```

RESULT 11

US-09-947-563-4

; Sequence 4, Application US/09947563

; Patent No. US20020156234A1

; GENERAL INFORMATION:

; APPLICANT: Rubroder, Franz-Josef

; Keller, Reinhold

; TITLE OF INVENTION: Improved process for obtaining

; insulin precursors having correctly bonded cystine
bridges

; NUMBER OF SEQUENCES: 7

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Finnegan, Henderson, Farrahaw, Garrett &

; Dunner

; STREET: 1300 I Street, N.W.

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20005-3315

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/947,563


```

; FILING DATE: 07-Sep-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/134,836
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Leslie McDonell
; REGISTRATION NUMBER: 34,872
; REFERENCE/DOCKET NUMBER: 02481.1600-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 96 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..96
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-947-563-4

```

```

Query Match          100.0%; Score 463; DB 9; Length 96;
Best Local Similarity 100.0%; Pred. No. 1.5e-44;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 70

Qy      61 SLQKRGIVEQCCTSICSLYQLENYCN 86
        ||||||||||||||||||||||||
Db      71 SLQKRGIVEQCCTSICSLYQLENYCN 96

```

RESULT 12

US-09-205-658-125

```

; Sequence 125, Application US/09205658
; Patent No. US20010029617A1
; GENERAL INFORMATION:
; APPLICANT: Ruvkun, Gary
; APPLICANT: Ogg, Scott
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
; TITLE OF INVENTION: IMPAIRED GLUCOSE TOLERANCE CONDITIONS
; FILE REFERENCE: 00786/351004
; CURRENT APPLICATION NUMBER: US/09/205,658
; CURRENT FILING DATE: 1998-12-03
; EARLIER APPLICATION NUMBER: 08/857,076
; EARLIER FILING DATE: 1997-05-15
; EARLIER APPLICATION NUMBER: 08/888,534
; EARLIER FILING DATE: 1997-07-07
; EARLIER APPLICATION NUMBER: US98/10080

```

; EARLIER FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 328
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 125
; LENGTH: 110
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-205-658-125

Query Match 100.0%; Score 463; DB 9; Length 110;
Best Local Similarity 100.0%; Pred. No. 1.8e-44;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
|
Db 25 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 84

Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 86
|
Db 85 SLQKRGIVEQCCTSICSLYQLENYCN 110

RESULT 13

US-09-815-229-3

; Sequence 3, Application US/09815229.
; Patent No. US20020058614A1
; GENERAL INFORMATION:
; APPLICANT: Filvaroff, Ellen H.
; APPLICANT: Okumu, Franklin W.
; TITLE OF INVENTION: USE OF INSULIN FOR THE TREATMENT OF CARTILAGENOUS
DISORDERS
; FILE REFERENCE: P1786R1US
; CURRENT APPLICATION NUMBER: US/09/815,229
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: US 60/192,103
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 3
; LENGTH: 110
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-815-229-3

Query Match 100.0%; Score 463; DB 9; Length 110;
Best Local Similarity 100.0%; Pred. No. 1.8e-44;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
|
Db 25 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 84

Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 86
|
Db 85 SLQKRGIVEQCCTSICSLYQLENYCN 110

RESULT 14

: PRIOR APPLICATION NUMBER: US 60/237,929

; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 110
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-969-748C-6

Query Match 100.0%; Score 463; DB 10; Length 110;
Best Local Similarity 100.0%; Pred. No. 1.8e-44;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
|
Db 25 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 84

Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 86
|
Db 85 SLQKRGIVEQCCTSICSLYQLENYCN 110

Search completed: February 11, 2005, 19:03:52
Job time : 63.5166 secs

OM protein - protein search, using sw model

Run on: February 11, 2005, 17:42:04 ; Search time 75.0517 Seconds
 (without alignments)
 586.780 Million cell updates/sec

Title: US-10-054-873-4
 Perfect score: 463
 Sequence: 1 FVNQHLCGSHLVEALYLVCG.....IVEQCCTSICSILYQLENYCN 86

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : UniProt_03:*
 1: uniprot_sprot:*
 2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

| Result | | | | | | | |
|--------|-------|-------|--------|----|------------|---------------------|--|
| No. | Score | Match | Length | DB | ID | Description | |
| 1 | 463 | 100.0 | 110 | 1 | INS_GORGO | Q6yk33 gorilla gor | |
| 2 | 463 | 100.0 | 110 | 1 | INS_HUMAN | P01308 homo sapien | |
| 3 | 463 | 100.0 | 110 | 1 | INS_PANTR | P30410 pan troglod | |
| 4 | 463 | 100.0 | 110 | 1 | INS_PONPY | Q8hxxv2 pongo pygma | |
| 5 | 456 | 98.5 | 110 | 1 | INS_CERAE | P30407 cercopithec | |
| 6 | 456 | 98.5 | 110 | 1 | INS_MACFA | P30406 macaca fasc | |
| 7 | 424 | 91.6 | 110 | 1 | INS_RABIT | P01311 oryctolagus | |
| 8 | 417 | 90.1 | 110 | 1 | INS_CANFA | P01321 canis famil | |
| 9 | 413 | 89.2 | 110 | 1 | INS_SPETR | Q91xi3 spermophilu | |
| 10 | 394 | 85.1 | 86 | 1 | INS_HORSE | P01310 equus cabal | |
| 11 | 394 | 85.1 | 110 | 1 | INS2_MOUSE | P01326 mus musculu | |
| 12 | 394 | 85.1 | 110 | 1 | INS2_RAT | P01323 rattus norv | |
| 13 | 392 | 84.7 | 108 | 1 | INS_AOTTR | P67972 aotus trivi | |
| 14 | 392 | 84.7 | 110 | 1 | INS_CRILLO | P01313 cricetulus | |
| 15 | 388 | 83.8 | 110 | 2 | Q8WNW6 | Q8wnw6 felis silve | |

| | | | | | | | |
|----|-------|------|-----|---|------------|--------|-------------|
| 16 | 385 | 83.2 | 110 | 1 | INS1_RAT | P01322 | rattus norv |
| 17 | 383 | 82.7 | 108 | 1 | INS_PIG | P01315 | sus scrofa |
| 18 | 377 | 81.4 | 110 | 1 | INS_PSAOB | Q62587 | psammomys o |
| 19 | 366.5 | 79.2 | 105 | 1 | INS_BOVIN | P01317 | bos taurus |
| 20 | 366 | 79.0 | 108 | 1 | INS1_MOUSE | P01325 | mus musculu |
| 21 | 362.5 | 78.3 | 105 | 1 | INS_SHEEP | P01318 | ovis aries |
| 22 | 342 | 73.9 | 65 | 2 | Q8HZ80 | Q8hz80 | pongo pygma |
| 23 | 342 | 73.9 | 65 | 2 | Q8HZ81 | Q8hz81 | gorilla gor |
| 24 | 334.5 | 72.2 | 108 | 1 | INS_RODSP | P21563 | rodentia sp |
| 25 | 314 | 67.8 | 110 | 1 | INS_CAVPO | P01329 | cavia porce |
| 26 | 277.5 | 59.9 | 109 | 1 | INS_OCTDE | P17715 | octodon deg |
| 27 | 276.5 | 59.7 | 103 | 1 | INS_SELRF | P51463 | selasphorus |
| 28 | 265.5 | 57.3 | 106 | 1 | INS2_XENLA | P12707 | xenopus lae |
| 29 | 265.5 | 57.3 | 107 | 1 | INS_CHICK | P67970 | gallus gall |
| 30 | 262.5 | 56.7 | 106 | 1 | INS1_XENLA | P12706 | xenopus lae |
| 31 | 256.5 | 55.4 | 51 | 1 | INS_BALPH | P67973 | balaenopter |
| 32 | 256.5 | 55.4 | 51 | 1 | INS_ELEMA | P01316 | elephas max |
| 33 | 256.5 | 55.4 | 51 | 1 | INS_PHYCA | P67974 | physeter ca |
| 34 | 256.5 | 55.4 | 81 | 1 | INS_ANAPL | P01333 | anas platyr |
| 35 | 256 | 55.3 | 96 | 2 | Q7MOU6 | Q7m0u6 | bacillus br |
| 36 | 254.5 | 55.0 | 51 | 2 | Q7MOG1 | Q7m0g1 | cricetidae |
| 37 | 251.5 | 54.3 | 51 | 1 | INS_ACOCA | P01324 | acomys cahi |
| 38 | 250.5 | 54.1 | 51 | 2 | Q7M217 | Q7m217 | canavalia e |
| 39 | 246.5 | 53.2 | 51 | 1 | INS_BALBO | P01314 | balaenopter |
| 40 | 246.5 | 53.2 | 51 | 1 | INS_CAMDR | P01320 | camelus dro |
| 41 | 246.5 | 53.2 | 51 | 1 | INS_CAPHI | P01319 | capra hircu |
| 42 | 246.5 | 53.2 | 106 | 2 | Q9I8Q7 | Q9i8q7 | rana pipien |
| 43 | 245.5 | 53.0 | 51 | 1 | INS_FELCA | P06306 | felis silve |
| 44 | 244.5 | 52.8 | 51 | 1 | INS_SAISC | P67971 | saimiri sci |
| 45 | 239.5 | 51.7 | 51 | 1 | INS_DIDMA | P18109 | didelphis m |

ALIGNMENTS

RESULT 1

INS_GORGO

ID INS_GORGO STANDARD; PRT; 110 AA.

AC Q6YK33;

DT 25-OCT-2004 (Rel. 45, Created)

DT 25-OCT-2004 (Rel. 45, Last sequence update)

DT 25-OCT-2004 (Rel. 45, Last annotation update)

DE Insulin precursor.

GN Name=INS;

OS Gorilla gorilla gorilla (Lowland gorilla).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.

OX NCBI_TaxID=9595;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=22833521; PubMed=12952878; DOI=10.1101/gr.948003;

RA Stead J.D.H., Hurles M.E., Jeffreys A.J.;

RT "Global haplotype diversity in the human insulin gene region.";

RL Genome Res. 13:2101-2111(2003).

CC -!- FUNCTION: Insulin decreases blood glucose concentration. It

CC increases cell permeability to monosaccharides, amino acids and

CC fatty acids. It accelerates glycolysis, the pentose phosphate

CC cycle, and glycogen synthesis in liver.
 CC -!- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
 CC disulfide bonds.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the insulin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AY137500; AAN06935.1; -.
 DR InterPro; IPR004825; Ins/IGF/relax.
 DR InterPro; IPR003234; Mollusc_ins.
 DR Pfam; PF00049; Insulin; 1.
 DR PRINTS; PR00277; INSULINB.
 DR ProDom; PD015667; Mollusc_ins; 1.
 DR SMART; SM00078; IIGF; 1.
 DR PROSITE; PS00262; INSULIN; 1.
 KW Glucose metabolism; Hormone; Insulin family; Signal.
 FT SIGNAL 1 24 By similarity.
 FT CHAIN 25 54 Insulin B chain.
 FT PROPEP 57 87 C peptide.
 FT CHAIN 90 110 Insulin A chain.
 FT DISULFID 31 96 Interchain (By similarity).
 FT DISULFID 43 109 Interchain (By similarity).
 FT DISULFID 95 100 By similarity.
 SQ SEQUENCE 110 AA; 11981 MW; C2C3B23B85E520E5 CRC64;

Query Match 100.0%; Score 463; DB 1; Length 110;
 Best Local Similarity 100.0%; Pred. No. 8e-41;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 25 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 84
 QY 61 SLQKRGIVEQCCTSICSLYQLENYCN 86
 ||||||||||||||||||||
 Db 85 SLQKRGIVEQCCTSICSLYQLENYCN 110

RESULT 2

INS_HUMAN

ID INS_HUMAN STANDARD; PRT; 110 AA.
 AC P01308;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Insulin precursor.
 GN Name=INS;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=80120725; PubMed=6243748;
 RA Bell G.I., Pictet R.L., Rutter W.J., Cordell B., Tischer E.,
 RA Goodman H.M.;
 RT "Sequence of the human insulin gene.";
 RL Nature 284:26-32(1980).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=80236313; PubMed=6248962;
 RA Ullrich A., Dull T.J., Gray A., Brosius J., Sures I.;
 RT "Genetic variation in the human insulin gene.";
 RL Science 209:612-615(1980).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=80054779; PubMed=503234;
 RA Bell G.I., Swain W.F., Pictet R.L., Cordell B., Goodman H.M.,
 RA Rutter W.J.;
 RT "Nucleotide sequence of a cDNA clone encoding human preproinsulin.";
 RL Nature 282:525-527(1979).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=80147417; PubMed=6927840;
 RA Sures I., Goeddel D.V., Gray A., Ullrich A.;
 RT "Nucleotide sequence of human preproinsulin complementary DNA.";
 RL Science 208:57-59(1980).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93364428; PubMed=8358440;
 RA Lucassen A.M., Bell J.I., Julier C., Lathrop M.;
 RT "Susceptibility to insulin dependent diabetes mellitus maps to a 4.1
 RT kb segment of DNA spanning the insulin gene and associated VNTR.";
 RL Nat. Genet. 4:305-310(1993).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [7]
 RP SEQUENCE OF 1-59 FROM N.A.
 RC TISSUE=Blood;
 RA Fajardy I.I., Weill J.J., Stuckens C.C., Danze P.M.P.;
 RT "Description of a novel RFLP diallelic polymorphism (~127 BsgI C/G)
 RT within the 5' region of insulin gene.";
 RL Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
 RN [8]
 RP SEQUENCE OF 25-54 AND 90-110.
 RX PubMed=14426955;
 RA Nicol D.S.H.W., Smith L.F.;
 RT "Amino-acid sequence of human insulin.";
 RL Nature 187:483-485(1960).
 RN [9]
 RP SEQUENCE OF 57-87.
 RX MEDLINE=71116410; PubMed=5101771;
 RA Oyer P.E., Cho S., Peterson J.D., Steiner D.F.;
 RT "Studies on human proinsulin. Isolation and amino acid sequence of the
 RT human pancreatic C-peptide.";
 RL J. Biol. Chem. 246:1375-1386(1971).
 RN [10]
 RP SEQUENCE OF 57-87.
 RX MEDLINE=71257722; PubMed=5560404;
 RA Ko A., Smyth D.G., Markussen J., Sundby F.;
 RT "The amino acid sequence of the C-peptide of human proinsulin.";
 RL Eur. J. Biochem. 20:190-199(1971).
 RN [11]
 RP SYNTHESIS.
 RX MEDLINE=75077277; PubMed=4443293;
 RA Sieber P., Kamber B., Hartmann A., Joehl A., Riniker B., Rittel W.;
 RT "Total synthesis of human insulin under directed formation of the
 RT disulfide bonds.";
 RL Helv. Chim. Acta 57:2617-2621(1974).
 RN [12]
 RP SYNTHESIS OF 57-87.
 RX MEDLINE=75040007; PubMed=4803504;
 RA Naithani V.K.;
 RT "Studies on polypeptides, IV. The synthesis of C-peptide of human
 RT proinsulin.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 354:659-672(1973).
 RN [13]
 RP SYNTHESIS OF 65-69 AND 70-73.
 RX MEDLINE=73161263; PubMed=4698555;
 RA Geiger R., Volk A.;
 RT "Synthesis of peptides with the properties of human proinsulin C
 RT peptides (hC peptide). 3. Synthesis of the sequences 14-17 and 9-13 of
 RT human proinsulin C peptides.";
 RL Chem. Ber. 106:199-205(1973).
 RN [14]
 RP SYNTHESIS OF 84-87.
 RX MEDLINE=73161261; PubMed=4698553;
 RA Geiger R., Jaeger G., Keonig W., Treuth G.;
 RT "Synthesis of peptides with the properties of human proinsulin C
 RT peptides (hC peptide). I. Scheme for the synthesis and preparation of
 RT the sequence 28-31 of human proinsulin C peptide.";
 RL Chem. Ber. 106:188-192(1973).

RN [15]
 RP VARIANT LOS ANGELES SER-48.
 RX MEDLINE=84016053; PubMed=6312455;
 RA Haneda M., Chan S.J., Kwok S.C.M., Rubenstein A.H., Steiner D.F.;
 RT "Studies on mutant human insulin genes: identification and sequence
 RT analysis of a gene encoding [SerB24]insulin.";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:6366-6370(1983).
 RN [16]
 RP VARIANTS LOS ANGELES SER-48 AND CHICAGO LEU-49.
 RX MEDLINE=84170233; PubMed=6424111;
 RA Shoelson S., Fickova M., Haneda M., Nahum A., Musso G., Kaiser E.T.,
 RA Rubenstein A.H., Tager H.;
 RT "Identification of a mutant human insulin predicted to contain a
 RT serine-for-phenylalanine substitution.";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:7390-7394(1983).
 RN [17]
 RP VARIANT PROVIDENCE ASP-34.
 RX MEDLINE=87175640; PubMed=3470784;
 RA Chan S.J., Seino S., Gruppuso P.A., Schwartz R., Steiner D.F.;
 RT "A mutation in the B chain coding region is associated with impaired
 RT proinsulin conversion in a family with hyperproinsulinemia.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:2194-2197(1987).
 RN [18]
 RP VARIANT WAKAYAMA LEU-92.
 RX MEDLINE=87058122; PubMed=3537011;
 RA Sakura H., Iwamoto Y., Sakamoto Y., Kuzuya T., Hirata H.;
 RT "Structurally abnormal insulin in a diabetic patient. Characterization
 RT of the mutant insulin A3 (Val-->Leu) isolated from the pancreas.";
 RL J. Clin. Invest. 78:1666-1672(1986).
 RN [19]
 RP VARIANT HIS-89.
 RX MEDLINE=90317021; PubMed=2196279;
 RA Barbetti F., Raben N., Kadowaki T., Cama A., Accili D., Gabbay K.H.,
 RA Merenich J.A., Taylor S.I., Roth J.;
 RT "Two unrelated patients with familial hyperproinsulinemia due to a
 RT mutation substituting histidine for arginine at position 65 in the
 RT proinsulin molecule: identification of the mutation by direct
 RT sequencing of genomic deoxyribonucleic acid amplified by polymerase
 RT chain reaction.";
 RL J. Clin. Endocrinol. Metab. 71:164-169(1990).
 RN [20]
 RP VARIANT HIS-89.
 RX MEDLINE=85261996; PubMed=4019786;
 RA Shibasaki Y., Kawakami T., Kanazawa Y., Akanuma Y., Takaku F.;
 RT "Posttranslational cleavage of proinsulin is blocked by a point
 RT mutation in familial hyperproinsulinemia.";
 RL J. Clin. Invest. 76:378-380(1985).
 RN [21]
 RP VARIANT KYOTO LEU-89.
 RX MEDLINE=92291307; PubMed=1601997;
 RA Yano H., Kitano N., Morimoto M., Polonsky K.S., Imura H., Seino Y.;
 RT "A novel point mutation in the human insulin gene giving rise to
 RT hyperproinsulinemia (proinsulin Kyoto).";
 RL J. Clin. Invest. 89:1902-1907(1992).
 RN [22]
 RP STRUCTURE BY NMR.
 RX MEDLINE=91104966; PubMed=2271664;

RA Hua Q.-X., Weiss M.A.;
 RT "Toward the solution structure of human insulin: sequential 2D 1H NMR
 RT assignment of a des-pentapeptide analogue and comparison with crystal
 RT structure.";
 RL Biochemistry 29:10545-10555(1990).
 RN [23]
 RP STRUCTURE BY NMR.
 RX MEDLINE=91242467; PubMed=2036420;
 RA Hua Q.-X., Weiss M.A.;
 RT "Comparative 2D NMR studies of human insulin and des-pentapeptide
 RT insulin: sequential resonance assignment and implications for protein
 RT dynamics and receptor recognition.";
 RL Biochemistry 30:5505-5515(1991).
 RN [24]
 RP STRUCTURE BY NMR.
 RX MEDLINE=91265527; PubMed=1646635; DOI=10.1016/0167-4838(91)90098-K;
 RA Hua Q.-X., Weiss M.A.;
 RT "Two-dimensional NMR studies of Des-(B26-B30)-insulin: sequence-
 RT specific resonance assignments and effects of solvent composition.";

Query Match 100.0%; Score 463; DB 1; Length 110;
 Best Local Similarity 100.0%; Pred. No. 8e-41;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 25 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 84

 Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 86
 ||||||||||||||||||||||||||||
 Db 85 SLQKRGIVEQCCTSICSLYQLENYCN 110

RESULT 3

INS_PANTR

ID INS_PANTR STANDARD; PRT; 110 AA.
 AC P30410;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Insulin precursor.
 GN Name=INS;
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 OX NCBI_TaxID=9598;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92219953; PubMed=1560757;
 RA Seino S., Bell G.I., Li W.;
 RT "Sequences of primate insulin genes support the hypothesis of a slower
 RT rate of molecular evolution in humans and apes than in monkeys.";
 RL Mol. Biol. Evol. 9:193-203(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22833521; PubMed=12952878; DOI=10.1101/gr.948003;
 RA Stead J.D.H., Hurles M.E., Jeffreys A.J.;

```

RT  "Global haplotype diversity in the human insulin gene region.";
RL  Genome Res. 13:2101-2111(2003).
CC  -!- FUNCTION: Insulin decreases blood glucose concentration. It
CC      increases cell permeability to monosaccharides, amino acids and
CC      fatty acids. It accelerates glycolysis, the pentose phosphate
CC      cycle, and glycogen synthesis in liver.
CC  -!- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
CC      disulfide bonds.
CC  -!- SUBCELLULAR LOCATION: Secreted.
CC  -!- SIMILARITY: Belongs to the insulin family.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; X61089; CAA43403.1; -.
DR  EMBL; AY137497; AAN06933.1; -.
DR  PIR; A42179; A42179.
DR  HSSP; P01308; 1AI0.
DR  InterPro; IPR004825; Ins/IGF/relax.
DR  Pfam; PF00049; Insulin; 1.
DR  PRINTS; PR00277; INSULINB.
DR  ProDom; PD015667; Mollusc_ins; 1.
DR  PROSITE; PS00262; INSULIN; 1.
KW  Glucose metabolism; Hormone; Insulin family; Signal.
FT  SIGNAL          1      24      By similarity.
FT  CHAIN           25      54      Insulin B chain.
FT  PROPEP          57      87      C peptide.
FT  CHAIN           90     110      Insulin A chain.
FT  DISULFID        31      96      Interchain (By similarity).
FT  DISULFID        43     109      Interchain (By similarity).
FT  DISULFID        95     100      By similarity.
SQ  SEQUENCE      110 AA;  12025 MW;  41EB8DF79837CEF5 CRC64;

```

```

Query Match          100.0%;  Score 463;  DB 1;  Length 110;
Best Local Similarity 100.0%;  Pred. No. 8e-41;
Matches   86;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

```

```

Qy      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 84

Qy      61 SLQKRGIVEQCCTSICSLYQLENYCN 86
        ||||||||||||||||||||||||||||
Db      85 SLQKRGIVEQCCTSICSLYQLENYCN 110

```

RESULT 4

INS_PONPY

```

ID  INS_PONPY          STANDARD;          PRT;   110 AA.
AC  Q8HXV2;
DT  05-JUL-2004 (Rel. 44, Created)
DT  05-JUL-2004 (Rel. 44, Last sequence update)

```

DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Insulin precursor.
 GN Name=INS;
 OS Pongo pygmaeus (Orangutan).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
 OX NCBI_TaxID=9600;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22833521; PubMed=12952878; DOI=10.1101/gr.948003;
 RA Stead J.D.H., Hurles M.E., Jeffreys A.J.;
 RT "Global haplotype diversity in the human insulin gene region.";
 RL Genome Res. 13:2101-2111(2003).
 CC -!- FUNCTION: Insulin decreases blood glucose concentration. It
 CC increases cell permeability to monosaccharides, amino acids and
 CC fatty acids. It accelerates glycolysis, the pentose phosphate
 CC cycle, and glycogen synthesis in liver.
 CC -!- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
 CC disulfide bonds.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the insulin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AY137503; AAN06937.1; -.
 DR HSSP; P01308; 1AI0.
 DR InterPro; IPR004825; Ins/IGF/relax.
 DR Pfam; PF00049; Insulin; 1.
 DR PRINTS; PR00277; INSULINB.
 DR ProDom; PD015667; Mollusc_ins; 1.
 DR SMART; SM00078; ILGF; 1.
 DR PROSITE; PS00262; INSULIN; 1.
 KW Glucose metabolism; Hormone; Insulin family; Signal.
 FT SIGNAL 1 24 By similarity.
 FT CHAIN 25 54 Insulin B chain.
 FT PROPEP 57 87 C peptide.
 FT CHAIN 90 110 Insulin A chain.
 FT DISULFID 31 96 Interchain (By similarity).
 FT DISULFID 43 109 Interchain (By similarity).
 FT DISULFID 95 100 By similarity.
 SQ SEQUENCE 110 AA; 12038 MW; 22D2B32B94F520F8 CRC64;

Query Match 100.0%; Score 463; DB 1; Length 110;
 Best Local Similarity 100.0%; Pred. No. 8e-41;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 25 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 84
 Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 86

|||||
Db 85 SLQKRGIVEQCCTSICSLYQLENYCN 110

RESULT 5

INS_CERAE

ID INS_CERAE STANDARD; PRT; 110 AA.
AC P30407; P01309;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Insulin precursor.
GN Name=INS;
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92219953; PubMed=1560757;
RA Seino S., Bell G.I., Li W.;
RT "Sequences of primate insulin genes support the hypothesis of a slower
RT rate of molecular evolution in humans and apes than in monkeys."
RL Mol. Biol. Evol. 9:193-203(1992).
RN [2]
RP SEQUENCE OF 57-87.
RX MEDLINE=72258016; PubMed=4626369;
RA Peterson J.D., Nehrlich S., Oyer P.E., Steiner D.F.;
RT "Determination of the amino acid sequence of the monkey, sheep, and
RT dog proinsulin C-peptides by a semi-micro Edman degradation
RT procedure."
RL J. Biol. Chem. 247:4866-4871(1972).
CC -!- FUNCTION: Insulin decreases blood glucose concentration. It
CC increases cell permeability to monosaccharides, amino acids and
CC fatty acids. It accelerates glycolysis, the pentose phosphate
CC cycle, and glycogen synthesis in liver.
CC -!- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
CC disulfide bonds.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the insulin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X61092; CAA43405.1; -.
DR PIR; B42179; B42179.
DR HSSP; P01308; 1AI0.
DR InterPro; IPR004825; Ins/IGF/relax.
DR Pfam; PF00049; Insulin; 1.
DR PRINTS; PR00277; INSULINB.
DR ProDom; PD015667; Mollusc_ins; 1.

DR SMART; SM00078; IIGF; 1.
 DR PROSITE; PS00262; INSULIN; 1.
 KW Direct protein sequencing; Glucose metabolism; Hormone;
 KW Insulin family; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 54 Insulin B chain.
 FT PROPEP 57 87 C peptide.
 FT CHAIN 90 110 Insulin A chain.
 FT DISULFID 31 96 Interchain.
 FT DISULFID 43 109 Interchain.
 FT DISULFID 95 100
 SQ SEQUENCE 110 AA; 12019 MW; 95A1F54BE7B247F9 CRC64;

Query Match 98.5%; Score 456; DB 1; Length 110;
 Best Local Similarity 98.8%; Pred. No. 4.3e-40;
 Matches 85; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
 |||
 Db 25 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDPQVGQVELGGGPGAGSLQPLALEG 84
 Qy 61 SLQKRGIIVEQCCTSICSLYQLENYCN 86
 |||
 Db 85 SLQKRGIIVEQCCTSICSLYQLENYCN 110

RESULT 6

INS_MACFA

ID INS_MACFA STANDARD; PRT; 110 AA.
 AC P30406; P01309;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Insulin precursor.
 GN Name=INS;
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
 OC Cercopithecinae; Macaca.
 OX NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83080474; PubMed=6184262; DOI=10.1016/0378-1119(82)90004-X;
 RA Wetekam W., Groneberg J., Leineweber M., Wengenmayer F.,
 RA Winnacker E.-L.;
 RT "The nucleotide sequence of cDNA coding for preproinsulin from the
 RT primate Macaca fascicularis."
 RL Gene 19:179-183(1982).
 CC -!- FUNCTION: Insulin decreases blood glucose concentration. It
 CC increases cell permeability to monosaccharides, amino acids and
 CC fatty acids. It accelerates glycolysis, the pentose phosphate
 CC cycle, and glycogen synthesis in liver.
 CC -!- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
 CC disulfide bonds.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the insulin family.
 CC -----

RA Devaskar S.U., Giddings S.J., Rajakumar P.A., Carnaghi L.R.,
 RA Menon R.K., Zahm D.S.;
 RT "Insulin gene expression and insulin synthesis in mammalian neuronal
 RT cells.";
 RL J. Biol. Chem. 269:8445-8454(1994).
 RN [2]
 RP SEQUENCE OF 25-54 AND 90-110.
 RX MEDLINE=66160119; PubMed=5949593; DOI=10.1016/0002-9343(66)90145-8;
 RA Smith L.F.;
 RT "Species variation in the amino acid sequence of insulin.";
 RL Am. J. Med. 40:662-666(1966).
 RN [3]
 RP SEQUENCE OF 56-110 FROM N.A.
 RA Giddings S.J., Carnaghi L.R., Devaskar S.U.;
 RL Submitted (APR-1991) to the EMBL/GenBank/DDBJ databases.
 CC -!- FUNCTION: Insulin decreases blood glucose concentration. It
 CC increases cell permeability to monosaccharides, amino acids and
 CC fatty acids. It accelerates glycolysis, the pentose phosphate
 CC cycle, and glycogen synthesis in liver.
 CC -!- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
 CC disulfide bonds.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the insulin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U03610; AAA19033.1; -.
 DR EMBL; M61153; AAA17540.1; -.
 DR PIR; A53438; INRB.
 DR HSSP; P01308; 1EV6.
 DR InterPro; IPR004825; Ins/IGF/relax.
 DR Pfam; PF00049; Insulin; 1.
 DR PRINTS; PR00277; INSULINB.
 DR ProDom; PD015667; Mollusc_ins; 1.
 DR SMART; SM00078; IIGF; 1.
 DR PROSITE; PS00262; INSULIN; 1.
 KW Direct protein sequencing; Glucose metabolism; Hormone;
 KW Insulin family; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 54 Insulin B chain.
 FT PROPEP 57 87 C peptide.
 FT CHAIN 90 110 Insulin A chain.
 FT DISULFID 31 96 Interchain.
 FT DISULFID 43 109 Interchain.
 FT DISULFID 95 100
 FT CONFLICT 83 83 E -> Y (in Ref. 3).
 SQ SEQUENCE 110 AA; 11838 MW; 82D2975B85D77FA8 CRC64;

 Query Match 91.6%; Score 424; DB 1; Length 110;
 Best Local Similarity 90.7%; Pred. No. 1e-36;
 Matches 78; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```

Qy      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
        |||||:||||| ||||| ||| |||
Db      25 FVNQHLCGSHLVEALYLVCGERGFFYTPKSRREVEELQVGQAEELGGGPGAGGLQPSALEL 84

Qy      61 SLQKRGIVEQCCTSICSLYQLENYCN 86
        :|||||
Db      85 ALQKRGIVEQCCTSICSLYQLENYCN 110

```

RESULT 8

INS_CANFA

```

ID      INS_CANFA          STANDARD;          PRT;    110 AA.
AC      P01321;
DT      21-JUL-1986 (Rel. 01, Created)
DT      21-JUL-1986 (Rel. 01, Last sequence update)
DT      05-JUL-2004 (Rel. 44, Last annotation update)
DE      Insulin precursor.
GN      Name=INS;
OS      Canis familiaris (Dog).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX      NCBI_TaxID=9615;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=83109071; PubMed=6296142;
RA      Kwok S.C.M., Chan S.J., Steiner D.F.;
RT      "Cloning and nucleotide sequence analysis of the dog insulin gene.
RT      Coded amino acid sequence of canine preproinsulin predicts an
RT      additional C-peptide fragment.";
RL      J. Biol. Chem. 258:2357-2363(1983).
RN      [2]
RP      SEQUENCE OF 25-54 AND 90-110.
RX      MEDLINE=66160119; PubMed=5949593; DOI=10.1016/0002-9343(66)90145-8;
RA      Smith L.F.;
RT      "Species variation in the amino acid sequence of insulin.";
RL      Am. J. Med. 40:662-666(1966).
CC      -!- FUNCTION: Insulin decreases blood glucose concentration. It
CC      increases cell permeability to monosaccharides, amino acids and
CC      fatty acids. It accelerates glycolysis, the pentose phosphate
CC      cycle, and glycogen synthesis in liver.
CC      -!- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
CC      disulfide bonds.
CC      -!- SUBCELLULAR LOCATION: Secreted.
CC      -!- SIMILARITY: Belongs to the insulin family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; V00179; CAA23475.1; -.
DR      PIR; A92413; IPDG.
DR      HSSP; P01317; 1APH.

```

DR InterPro; IPR004825; Ins/IGF/relax.
 DR Pfam; PF00049; Insulin; 1.
 DR PRINTS; PR00277; INSULINB.
 DR ProDom; PD015667; Mollusc_ins; 1.
 DR SMART; SM00078; IIGF; 1.
 DR PROSITE; PS00262; INSULIN; 1.
 KW Direct protein sequencing; Glucose metabolism; Hormone;
 KW Insulin family; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 54 Insulin B chain.
 FT PROPEP 57 87 C peptide.
 FT CHAIN 90 110 Insulin A chain.
 FT DISULFID 31 96 Interchain.
 FT DISULFID 43 109 Interchain.
 FT DISULFID 95 100
 SQ SEQUENCE 110 AA; 12190 MW; A574791864A4FB98 CRC64;

Query Match 90.1%; Score 417; DB 1; Length 110;
 Best Local Similarity 89.5%; Pred. No. 5.4e-36;
 Matches 77; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
 |||||
 Db 25 FVNQHLCGSHLVEALYLVCGERGFFYTPKARREVEDLQVRDELAPGEGGLQPLALEG 84
 Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 86
 :|||
 Db 85 ALQKRGIVEQCCTSICSLYQLENYCN 110

RESULT 9

INS_SPETR

ID INS_SPETR STANDARD; PRT; 110 AA.
 AC Q91XI3;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Insulin precursor.
 GN Name=INS;
 OS *Spermophilus tridecemlineatus* (Thirteen-lined ground squirrel).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Sciurinae;
 OC *Spermophilus*.
 OX NCBI_TaxID=43179;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RA Tredrea M.M., Buck M.J., Guhaniyogi J., Squire T.L., Andrews M.T.;
 RT "Regulation of PDK4 expression in a hibernating mammal."
 RL Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
 CC -!- FUNCTION: Insulin decreases blood glucose concentration. It
 CC increases cell permeability to monosaccharides, amino acids and
 CC fatty acids. It accelerates glycolysis, the pentose phosphate
 CC cycle, and glycogen synthesis in liver.
 CC -!- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
 CC disulfide bonds.
 CC -!- SUBCELLULAR LOCATION: Secreted.

```

CC      -!- SIMILARITY: Belongs to the insulin family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AY038604; AAK72558.1; -.
DR      HSSP; P01308; 1EV6.
DR      InterPro; IPR004825; Ins/IGF/relax.
DR      Pfam; PF00049; Insulin; 1.
DR      PRINTS; PR00277; INSULINB.
DR      ProDom; PD015667; Mollusc_ins; 1.
DR      SMART; SM00078; IIGF; 1.
DR      PROSITE; PS00262; INSULIN; 1.
KW      Glucose metabolism; Hormone; Insulin family; Signal.
FT      SIGNAL          1      24      By similarity.
FT      CHAIN           25      54      Insulin B chain.
FT      PROPEP          57      87      C peptide.
FT      CHAIN           90     110      Insulin A chain.
FT      DISULFID        31      96      Interchain (By similarity).
FT      DISULFID        43     109      Interchain (By similarity).
FT      DISULFID        95     100      By similarity.
SQ      SEQUENCE      110 AA;  12004 MW;  4511768D6622BEE5 CRC64;

```

```

Query Match          89.2%;  Score 413;  DB 1;  Length 110;
Best Local Similarity 89.5%;  Pred. No. 1.4e-35;
Matches    77;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

```

```

Qy      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
        |||
Db      25 FVNQHLCGSHLVEALYLVCGERGFFYTPKSRREVVEEQGGQVELGGGPGAGLPQPLALEM 84
        |||
Qy      61 SLQKRGIVEQCCTSICSLYQLENYCN 86
        :|||
Db      85 ALQKRGIVEQCCTSICSLYQLENYCN 110

```

RESULT 10

INS_HORSE

```

ID      INS_HORSE      STANDARD;      PRT;      86 AA.
AC      P01310;
DT      21-JUL-1986 (Rel. 01, Created)
DT      21-JUL-1986 (Rel. 01, Last sequence update)
DT      25-OCT-2004 (Rel. 45, Last annotation update)
DE      Insulin precursor.
GN      Name=INS;
OS      Equus caballus (Horse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX      NCBI_TaxID=9796;
RN      [1]
RP      SEQUENCE OF 1-30 AND 66-86.
RX      PubMed=13373434;

```


DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Insulin 2 precursor.
 GN Name=Ins2; Synonyms=Ins-2;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87169768; PubMed=3104603;
 RA Wentworth B.M., Schaefer I.M., Villa-Komaroff L., Chirgwin J.M.;
 RT "Characterization of the two nonallelic genes encoding mouse
 RT preproinsulin.";
 RL J. Mol. Evol. 23:305-312(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NON;
 RX MEDLINE=90372989; PubMed=2397023;
 RA Sawa T., Ohgaku S., Morioka H., Yano S.;
 RT "Molecular cloning and DNA sequence analysis of preproinsulin genes in
 RT the NON mouse, an animal model of human non-obese, non-insulin-
 RT dependent diabetes mellitus.";
 RL J. Mol. Endocrinol. 5:61-67(1990).
 RN [3]
 RP SEQUENCE OF 25-54 AND 90-110.
 RX MEDLINE=72189455; PubMed=5063718;
 RA Buenzli H.F., Glatthaar B., Kunz P., Muelhaupt E., Humbel R.E.;
 RT "Amino acid sequence of the two insulins from mouse (Maus musculus).";
 RL Hoppe-Seyler's Z. Physiol. Chem. 353:451-458(1972).
 CC -!- FUNCTION: Insulin decreases blood glucose concentration. It
 CC increases cell permeability to monosaccharides, amino acids and
 CC fatty acids. It accelerates glycolysis, the pentose phosphate
 CC cycle, and glycogen synthesis in liver.
 CC -!- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
 CC disulfide bonds.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the insulin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X04724; CAA28433.1; -.
 DR PIR; A26342; INMS2.
 DR HSSP; P01317; 1APH.
 DR MGD; MGI:96573; Ins2.
 DR GO; GO:0005634; C:nucleus; IDA.
 DR GO; GO:0005732; C:small nucleolar ribonucleoprotein complex; IDA.
 DR GO; GO:0000187; P:activation of MAPK; IDA.
 DR GO; GO:0006006; P:glucose metabolism; IMP.
 DR GO; GO:0008286; P:insulin receptor signaling pathway; IDA.
 DR GO; GO:0016042; P:lipid catabolism; IDA.
 DR GO; GO:0042981; P:regulation of apoptosis; IMP.

DR GO; GO:0042325; P:regulation of phosphorylation; IDA.
 DR GO; GO:0006983; P:response to ER-overload; IMP.
 DR InterPro; IPR004825; Ins/IGF/relax.
 DR Pfam; PF00049; Insulin; 1.
 DR PRINTS; PR00277; INSULINB.
 DR ProDom; PD015667; Mollusc_ins; 1.
 DR SMART; SM00078; IIGF; 1.
 DR PROSITE; PS00262; INSULIN; 1.
 KW Direct protein sequencing; Glucose metabolism; Hormone;
 KW Insulin family; Multigene family; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 54 Insulin 2 B chain.
 FT PROPEP 57 87 C peptide.
 FT CHAIN 90 110 Insulin 2 A chain.
 FT DISULFID 31 96 Interchain.
 FT DISULFID 43 109 Interchain.
 FT DISULFID 95 100
 SQ SEQUENCE 110 AA; 12364 MW; 3554C8803D24FDAD CRC64;

Query Match 85.1%; Score 394; DB 1; Length 110;
 Best Local Similarity 84.9%; Pred. No. 1.4e-33;
 Matches 73; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
 || ||||| :|| || || :||| || || || ||
 Db 25 FVKQHLCGSHLVEALYLVCGERGFFYTPMSRREVEDPQVAQLLEGGGPGAGDLQTLALEV 84
 Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 86
 : ||||| :||| ||||| |||||
 Db 85 AQQKRGIVDQCCTSICSLYQLENYCN 110

RESULT 12

INS2_RAT

ID INS2_RAT STANDARD; PRT; 110 AA.
 AC P01323;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Insulin 2 precursor.
 GN Name=Ins2; Synonyms=Ins-2;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Liver;
 RX MEDLINE=80045035; PubMed=498284; DOI=10.1016/0092-8674(79)90071-0;
 RA Lomedico P., Rosenthal N., Efstratiadis A., Gilbert W., Kolodner R.,
 RA Tizard R.;
 RT "The structure and evolution of the two nonallelic rat preproinsulin
 RT genes.";
 RL Cell 18:545-558(1979).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86310882; PubMed=2427930;

RA Soares M.B., Schin E., Henderson A., Karathanasis S.K., Cate R.,
 RA Zeitlin S., Chirgwin J., Efstratiadis A.;
 RT "RNA-mediated gene duplication: the rat preproinsulin I gene is a
 RT functional retroposon.";
 RL Mol. Cell. Biol. 5:2090-2103(1985).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=80240379; PubMed=6249167;
 RA Lomedico P.T., Rosenthal N., Kolodner R., Efstratiadis A., Gilbert W.;
 RT "The structure of rat preproinsulin genes.";
 RL Ann. N. Y. Acad. Sci. 343:425-432(1980).
 RN [4]
 RP SEQUENCE OF 25-54 AND 90-110.
 RX MEDLINE=70067613; PubMed=4311938;
 RA Steiner D.F., Clark J.L., Nolan C., Rubenstein A.H., Margoliash E.,
 RA Aten B., Oyer P.E.;
 RT "Proinsulin and the biosynthesis of insulin.";
 RL Recent Prog. Horm. Res. 25:207-282(1969).
 RN [5]
 RP SEQUENCE OF 57-87.
 RX MEDLINE=73061498; PubMed=4640931;
 RA Tager H.S., Steiner D.F.;
 RT "Primary structures of the proinsulin connecting peptides of the rat
 RT and the horse.";
 RL J. Biol. Chem. 247:7936-7940(1972).
 RN [6]
 RP SEQUENCE OF 57-87, AND REVISIONS.
 RX MEDLINE=72177385; PubMed=4554104;
 RA Markussen J., Sundby F.;
 RT "Rat-proinsulin C-peptides. Amino-acid sequences.";
 RL Eur. J. Biochem. 25:153-162(1972).
 CC -!- FUNCTION: Insulin decreases blood glucose concentration. It
 CC increases cell permeability to monosaccharides, amino acids and
 CC fatty acids. It accelerates glycolysis, the pentose phosphate
 CC cycle, and glycogen synthesis in liver.
 CC -!- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
 CC disulfide bonds.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the insulin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; V01243; CAA24560.1; -.
 DR EMBL; J00748; AAA41443.1; -.
 DR EMBL; M25585; AAA41440.1; -.
 DR EMBL; M25583; AAA41440.1; JOINED.
 DR PIR; B90789; IPRT2.
 DR HSSP; P01317; 1APH.
 DR RGD; 2916; Ins2.
 DR InterPro; IPR004825; Ins/IGF/relax.
 DR Pfam; PF00049; Insulin; 1.

DR PRINTS; PR00277; INSULINB.
DR ProDom; PD015667; Mollusc_ins; 1.
DR SMART; SM00078; IIGF; 1.
DR PROSITE; PS00262; INSULIN; 1.
KW Direct protein sequencing; Glucose metabolism; Hormone;
KW Insulin family; Multigene family; Signal.
FT SIGNAL 1 24
FT CHAIN 25 54 Insulin 2 B chain.
FT PROPEP 57 87 C peptide.
FT CHAIN 90 110 Insulin 2 A chain.
FT DISULFID 31 96 Interchain.
FT DISULFID 43 109 Interchain.
FT DISULFID 95 100
SQ SEQUENCE 110 AA; 12339 MW; 3A626DA98C86F3CA CRC64;

Query Match 85.1%; Score 394; DB 1; Length 110;
Best Local Similarity 84.9%; Pred. No. 1.4e-33;
Matches 73; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
|| ||||| :||| || || :||| || || |||||
Db 25 FVKQHLGSHLVEALYLVCGERGFFYTPMSRREVEDPQVAQLELGGGPGAGDLQTLALEV 84
Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 86
: |||||:|||||
Db 85 ARQKRGIVDQCCTSICSLYQLENYCN 110

RESULT 13

INS_AOTTR

ID INS_AOTTR STANDARD; PRT; 108 AA.
AC P67972; P10604;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Insulin precursor.
GN Name=INS;
OS Aotus trivirgatus (Night monkey) (Douroucouli).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Aotinae; Aotus.
OX NCBI_TaxID=9505;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88041119; PubMed=3118367;
RA Seino S., Steiner D.F., Bell G.I.;
RT "Sequence of a New World primate insulin having low biological potency
and immunoreactivity.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:7423-7427(1987).
CC -!- FUNCTION: Insulin decreases blood glucose concentration. It
increases cell permeability to monosaccharides, amino acids and
fatty acids. It accelerates glycolysis, the pentose phosphate
cycle, and glycogen synthesis in liver.
CC -!- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
disulfide bonds.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the insulin family.
CC -----


```

RA Bell G.I., Sanchez-Pescador R.;
RT "Sequence of a cDNA encoding Syrian hamster preproinsulin.";
RL Diabetes 33:297-300(1984).
RN [2]
RP SEQUENCE OF 25-54 AND 90-110.
RA Neelon F.A., Delcher H.K., Steinman H., Lebovitz H.E.;
RT "Structure of hamster insulin: comparison with a tumor insulin.";
RL Fed. Proc. 32:300-300(1973).
CC -!- FUNCTION: Insulin decreases blood glucose concentration. It
CC increases cell permeability to monosaccharides, amino acids and
CC fatty acids. It accelerates glycolysis, the pentose phosphate
CC cycle, and glycogen synthesis in liver.
CC -!- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
CC disulfide bonds.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the insulin family.

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

```

DR      EMBL; M26328; AAA37089.1; -.
DR      HSSP; P01308; 1EV6.
DR      InterPro; IPR004825; Ins/IGF/relax.
DR      Pfam; PF00049; Insulin; 1.
DR      PRINTS; PR00277; INSULINB.
DR      ProDom; PD015667; Mollusc_ins; 1.
DR      SMART; SM00078; ILGF; 1.
DR      PROSITE; PS00262; INSULIN; 1.
KW      Direct protein sequencing; Glucose metabolism; Hormone;
KW      Insulin family; Signal.
FT      SIGNAL          1          24
FT      CHAIN           25         54      Insulin B chain.
FT      PROPEP          57         87      C peptide.
FT      CHAIN           90        110      Insulin A chain.
FT      DISULFID         31         96      Interchain.
FT      DISULFID         43        109      Interchain.
FT      DISULFID         95        100
SO      SEQUENCE        110 AA;  12268 MW;  219E92B85A535CEC CRC64;

```

Query Match 84.7%; Score 392; DB 1; Length 110;
Best Local Similarity 84.9%; Pred. No. 2.3e-33;
Matches 73; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

```
Qy      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
          ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     25 FVNQHLCGSHLVEALYLVCGERGFFYTPKSRRGVEDPQVAQLLEGGGPGADDLQTLEAV 84

Qy     61 SLQKRGI VEQCCTSICSLYQLENYCN 86
          : ||||| : ||||| ||||| ||||| |||||
Db    85 AQQKRGI VDQCCTSICSLYQLENYCN 110
```

Q8WNW6

Query Match 83.8%; Score 388; DB 2; Length 110;
Best Local Similarity 83.7%; Pred. No. 6.1e-33;
Matches 72; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

```
Qy      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
        |||||||
Db     25 FVNQHLCGSHLVEALYLVCGERGFFYTPKARREAEDLQGKDAELGEAPGAGGLQPSALEA 84

Qy     61 SLQKRGIIVEQCCTSICSLYQLENYCN 86
        |||||
Db    85 PLQKRGIIVEQCCASVCSLYQLEHYCN 110
```

Search completed: February 11, 2005, 18:22:48
Job time : 76.0517 secs